Systems level generation of mammalian circadian rhythms and its connection to liver metabolism

DISSERTATION

zur Erlangung des akademischen Grades

Doctor rerum naturalium
(Dr. rer. nat.)

eingereicht an der
Lebenswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin

von

M.Sc., Jan Patrick, PETT

Präsidentin
der Humboldt-Universität zu Berlin:

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin:

Prof. Dr. Bernhard Grimm

Gutachter/innen:

1. Prof. Dr. Hanspeter Herzel
2. Prof. Dr. Nils Blüthgen
3. Prof. Dr. Didier Gonze

Tag der mündlichen Prüfung: 30.4.2019

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Abstract

Lebenswissenschaftlichen Fakultät
Institute for Theoretical Biology

Doctor rerum naturalium

Systems level generation of mammalian circadian rhythms and its connection to liver metabolism

by Jan Patrick, PETT

Circadian clocks are endogenous oscillators that generate 24-hour rhythms. They allow many organisms to synchronize their physiology and behaviour with daily changes of the environment. In mammals such clocks are based on transcriptional-translational feedbacks and exist in almost every cell. Within an organism different clocks are distinguished by their localization in different organs, which is connected with differences in the functions they modulate in a daily manner. One of the key physiological functions of circadian clocks in various organs seems to be the temporal alignment of metabolic processes.

The general mechanism of feedback regulation that generates circadian rhythmicity has been uncovered, but it is not fully understood which feedback loops contribute to rhythm generation. A group of so called core clock genes that appear to be relevant for generating rhythmicity is connected in a complex network that comprises a multitude of such loops. In principal, a single negative feedback loop can be sufficient, but additional loops may help to generate rhythms.

In the first publication contained in this thesis we introduced and applied a method to systematically test regulations in a data-driven mathematical model of the core clock. Surprisingly, we discovered a type of feedback loop that has previously not been considered in the context of the mammalian circadian clock. This repressilator is consistent with knockout studies and further perturbation experiments and could constitute an explanation for different phases observed between cryptochromes, that are part of the core clock. Since the repressilator connects other loops that have been described before, we suggested that it could generate rhythms in synergy with these coherent motifs.

In the second publication presented here, we analyzed the core clock regarding a possible tissue-specificity. Given the different functions of organs and multiple co-existing feedback loops in the network it could be possible that different loops generate rhythms in different tissues. To this end, we repeatedly fitted the same mathematical model to tissue-specific data sets and obtained ensembles of model versions. With an adjusted version of the method introduced in the first part of

https://doi.org/10.18452/19960

the work we identified essential feedback loops in all model versions. Interestingly, for all tissue-specific data sets we found synergies of loops generating rhythms together, as previously suggested. Further, we found that the synergies which generate rhythms differ depending on the tissue, with the most marked differences between central clocks in the brain and peripheral clocks in other parts of the body. These differences could be connected to different physiological functions exerted by the clocks in various organs as well as differences in their routes of synchronization.

In the third part of the work we examined the circadian aspects of metabolism, which seems to be an important function in particular of peripheral clocks. A large number of circadian omics studies is available in the literature for mouse liver. We identified rhythmic components in different types of such studies, integrated and mapped them to a large-scale metabolic network. Our analysis confirmed that many metabolic pathways may follow circadian rhythms. Interestingly, we found that the average peak times of rhythmic components between various pathways differ. Such differences might reflect a temporal alignment of metabolic functions to the time when they are required and could be due to the association of different core clock components with pathways.

In summary, the core clock network was analyzed and the mechanisms underlying circadian rhythm generation examined, which led us to the identification of a newly described feedback loop. Further, we found that the mechanisms that generate these rhythms might differ between tissues, in particular between peripheral and brain tissues which possibly reflects different functions. To characterize the circadian function and connections between organs, we performed a large-scale analysis of daily variations in metabolism—a particularly important physiological aspect of peripheral tissues—and found a temporal structure of circadian regulation. The generation of circadian rhythms under different conditions and their connection to the daytime-dependent metabolic cycles are key questions to understand the organ-specific function of circadian clocks and could have an impact on the treatment of metabolic disease.

https://doi.org/10.18452/19960

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Zusammenfassung

Lebenswissenschaftlichen Fakultät
Institute for Theoretical Biology

Doctor rerum naturalium

Systems level generation of mammalian circadian rhythms and its connection to liver metabolism

von Jan Patrick, PETT

Circadiane Uhren sind endogene Oszillatoren, die 24-Stunden Rhythmen generieren. Sie erlauben vielen Organismen deren Physiologie und Verhalten an tägliche Änderungen der Umwelt anzupassen. In Säugetieren basieren solche Uhren auf transkriptional-translationalen Rückkopplungsschleifen und existieren in fast jeder Zelle. Innerhalb eines Organismus werden verschiedene Uhren durch ihre Lage in Organen unterschieden, die mit verschiedenen Funktionen verknüpft sind, welche sie in einer täglichen Weise modulieren. Eine der physiologischen Schlüsselfunktionen von cirkadianen Uhren in verschiedenen Organen scheint die zeitliche Anordnung von metabolischen Prozessen zu sein.

Der generelle Mechanismus einer Rückkopplung, die cirkadiane Rhythmen erzeugt wurde bereits entdeckt, aber es ist noch nicht ganz verstanden, welche Rückkopplungsschleifen zur Erzeugung von Rhythmen beitragen. Eine Gruppe so genannter Kernuhr-Gene, die für die Erzeugung von Rhythmizität relevant sind, ist in einem komplizierten Netzwerk verbunden, das eine Vielzahl solcher Schleifen enthält. Im Prinzip kann eine einzelne negative Rückkopplungsschleife ausreichen, doch zusätzliche Schleifen können helfen Rhythmen zu erzeugen.

In der ersten Publikation, die in dieser Dissertation enthalten ist, haben wir eine Methode eingeführt and angewandt, um systematisch Regulationen in einem datengetriebenen mathematischen Modell der Kernuhr zu testen. Überraschenderweise haben wir einen Typ von Rückkopplungsschleife entdeckt, der vorher noch nicht im Zusammenhang mit der circadianen Uhr in Säugetieren in Betracht gezogen wurde. Dieser Repressilator ist mit Gen-knockout Studien und weiteren Perturbationsexperimenten konsistent und könnte eine Erklärung für verschiedene Phasen zwischen Chryptochromen liefern, die ein Teil der Kernuhr sind. Da der Repressilator andere Schleifen verbindet, die zuvor beschrieben wurden, haben wir vorgeschlagen, dass er Rhythmen in Synergie mit diesen kohärenten Motiven erzeugen könnte.

In der zweiten Publikation, die hier vorgestellt wird, haben wir die Kernuhr hinsichtlich einer möglichen Gewebespezifität untersucht. Hinsichtlich der verschiedenen Funktionen von Organen und vielen co-existierenden Rückkopplungsschleifen

https://doi.org/10.18452/19960

im Netzwerk könnte es möglich sein, dass in verschiedenen Geweben unterschiedliche Rückkopplungschleifen Rhythmen erzeugen. Daher haben wir das selbe Modell wiederholt an gewebespezifische Datensätze angepasst, um Ensembles von Modellversionen zu erzeugen. Mit einer modifizierten Version der Methode, die im ersten Teil der Arbeit eingeführt wurde, haben wir essentielle Rückkopplungsschleifen in allen Modellversionen identifiziert. Interessanterweise fanden wir dabei für alle gewebespezifischen Datensätze Synergien von Rückkopplungsschleifen, die zusammen Rhythmen erzeugen, wie zuvor in Betracht gezogen. Desweiteren haben wir festgestellt, dass die Synergien, die Rhythmen erzeugen, sich abhängig vom Gewebe unterscheiden – mit den meisten Unterschieden zwischen zentralen Uhren im Gehirn und peripheren Uhren in anderen Organen des Körpers. Diese Unterschiede könnten mit verschiedenen physiologischen Funktionen der Uhren in verschiedenen Organen sowie Unterschieden zwischen deren Wegen der Synchronisation zusammenhängen.

In dem dritten Teil der Arbeit haben wir die circadianen Aspekte des Metabolismus untersucht, der eine wichtige Funktion insbesondere der peripheren Uhren zu sein scheint. Eine große Zahl an circadianen Omics-Studien der Mausleber ist in der wissenschaftlichen Literatur verfügbar. Wir haben circadiane Komponenten in verschiedenen Typen solcher Studien identifiziert, integriert und auf ein großskaliges metabolisches Netzwerk abgebildet. Unsere Analyse hat bestätigt, dass viele Stoffwechselwege vermutlich circadianen Rhythmen folgen. Interessanterweise haben wir festgestellt, dass die durchschnittlichen Phasen von rhythmischen Komponenten sich zwischen verschiedenen Stoffwechselwegen unterscheiden. Solche Unterschiede könnten eine zeitliche Anpassung metabolischer Funktionen an die Zeiten darstellen zu denen sie gebraucht werden und durch die Assoziation verschiedener Kernuhrelemente mit Stoffwechselwegen bedingt sein.

Zusammengefasst, haben wir das Kernuhr-Netzwerk analysiert und die der Erzeugung von circadianen Rhythmen zugrungeliegenden Mechanismen untersucht, wodurch wir eine neu beschriebene Rückkopplungsschleife identifiziert haben. Desweiteren haben wir herausgefunden, dass die Mechanismen, die diese Rhythmen erzeugen, sich zwischen Geweben unterscheiden könnten, insbesondere zwischen peripheren- und Gehirngeweben, was möglicherweise verschiedene Funktionen wiederspiegelt. Um die circadianen Funktionen und Verbindungen zwischen Organen zu charakterisieren, haben wir eine großskalige Analyse der täglichen Variationen des Stoffwechsels – eines besonders wichtigen physiologischen Aspekts peripherer Gewebe – durchgeführt und eine zeitliche Strukturierung der circadianen Regulation vorgefunden. Die Erzeugung circadianer Rhythmen unter verschiedenen Bedingungen und ihre Verbindung zu tageszeitabhängigen metabolischen Zyklen sind Kernfragen zum Verständnis der organspezifischen Funktion circadianer Uhren und könnten einen Einfluss auf die Behandlung von metabolischen Krankheiten haben.

https://doi.org/10.18452/19960

Acknowledgements

First of all I want to thank my supervisor Hanspeter Herzel. He supported me in all possible ways and consulted me on the scientific aspect of my work, while at the same time giving me the freedom to follow my own ideas. He was always ready to assist with problems and encouraged me to attend many conferences and symposia, which remain as unique memories.

Further, I want to thank my second supervisor Nils Blüthgen and TAC member Achim Kramer for their consultations during meetings and retreats and Matthias König for supervising me in the third unpublished project described in this thesis.

I also want thank my colleagues Christoph Schmal, Bharath Ananthasubramaniam, Grisha Bordyugov, Abhishek Uphadyay and Marta del Olmo for discussions and inspiration, besides contributing to a perfect working environment.

I am further grateful to Gen Kurosawa, Isao Tokuda and Jihwan Myung for inviting me to stay in their labs and give presentations as well as consulting me on topics of my work.

Finally, I want to give a special thanks my family and my wife for their loving support over the last years.

https://doi.org/10.18452/19960

Declaration of Authorship

I, Jan Patrick, PETT, declare that this thesis titled, "Systems level generation of mammalian circadian rhythms and its connection to liver metabolism" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given.

Signed:

Date:

https://doi.org/10.18452/19960

Contents

https://doi.org/10.18452/19960

List of Figures

https://doi.org/10.18452/19960

Chapter 1

Introduction

1.1 Biological Background

While a separate introduction is given in sections 2.1.1, 2.2.1 and 2.3.1 for each part of the work presented in this thesis, this section contains a brief overview combining aspects of all three parts.

1.1.1 Circadian clocks, their inner workings and role in physiology

Circadian clocks are endogenous timing systems that generate 24-hour rhythms synchronized with the day-night cycle. They have been evolved by many organisms to anticipate environmental changes throughout the day and to adapt their physiology and behaviour accordingly [1]. The mechanisms that generate these autonomous rhythms in mammals are present in almost every cell. So called *core clock genes*, connected in a gene transcriptional-translational network, create oscillations of their own transcript and protein levels [2]. These cellular oscillators synchronize their rhythms with each other and with daily environmental changes via a large diversity of organ-dependent signals [2, 3].

From theory it is known, that feedback loops are necessary to generate rhythms [4, 5]. In case of the mammalian circadian clock *transcriptional-translational feedback loops* (TTFL) over the core clock genes are considered as the key mechanism [2, 6, 7]. Starting from the discovery of a fist loop [8]—found in *Drosophila* but with homologous genes in mammals—a growing number of components has been discovered that form additional loops [9]. The benefit of multiple connected feedback loops for rhythm generation has been described from a theoretical perspective [5, 10]. Further, in particular one loop over nuclear receptors that are part of the core clock network [11] has been considered for mammalian circadian rhythm generation [2, 12]. However, in principal many additional loops exist that could contribute to the generation of circadian oscillation in mammals [9].

Given tissue-dependent differences, a variety of *central* and *peripheral clocks* are distinguished, which are situated in various organs of the body [3]. Clocks in the brain are typically called central, with the *suprachiasmatic nucleus* (SCN), that receives light signals directly via the eye nerves, as a key component [2, 13]. The SCN

https://doi.org/10.18452/19960

synchronizes peripheral clocks in other organs via various signals [14–16]. However, there are multiple routes of synchronization signals between all clocks of the body [3]. For peripheral clocks in particular the feeding-time and connected variations of metabolites in the blood stream seem to play an important role [17, 18]. While the time of feeding can be seen as a behaviorally controlled signal from the SCN, metabolite variations are varied by rhythmic uptake and release from different organs. The liver, as one of the biggest organs, has been studied extensively in chronobiology and exhibits strong rhythms [19]. It receives nutrient rich blood directly from the portal vein and participates in numerous metabolic functions.

The transcription factors comprising the core clock transmit rhythms to a large amount of other genes and proteins to control physiology in a circadian manner [20, 21]. However, the sets of such rhythmically expressed *clock controlled genes* (CCGs) vary from organ to organ [20, 22, 23]. In particular, many CCGs point to a function in metabolism probably connected to the rhythmic processing of nutrients and varying energy demands [20, 24, 25]. Thus, metabolism has been recognized as an important aspect of circadian physiology [26]. While the circadian clock exerts effect on metabolism directly or via CCGs, concentrations of metabolites also entrain the rhythms of the clock [19], resulting in a complex mutual relationship.

Temporal alignment or separation of physiological processes is seen as an important function of the clock, as disruption of circadian rhythms has been associated with various diseases including sleep disorders [27, 28], some types of cancer [29, 30] as well as obesity, diabetes and metabolic syndrome [31–33]. Further, elevated blood glucose and insulin concentrations as well as diverse metabolic disorders were shown to be associated to circadian misalignments in otherwise healthy persons [34–36]. Therefore, understanding rhythms generated by the circadian clock and their connection to physiology may as well benefit the understanding of diseases and their treatment.

1.1.2 Motivation and project outline

Since the core clock network is formed by multiple feedbacks [9], the question arises which loops contribute to the generation of rhythms and are therefore part of a core clock mechanism. Most experimental studies that examine the effect of perturbations target genes. However, one gene may be a part of several feedback loops, which makes a detailed analysis difficult. We designed an *in silico* study in which we analyzed a mathematical model [23] that reflects the complexity of feedback mechanisms and tested loops systematically. Interestingly, we found that a negative feedback loop is responsible for the generation of rhythms that has not been described before in the mammalian circadian clock. A comparison with the literature showed that the mechanism we found is consistent with knockout studies and some more rarely performed constitutive expression experiments. We suggest the use of such constitutive expression experiments to further validate our findings.

https://doi.org/10.18452/19960

Given the differences between organs, a second question concerns a possible tissue-specificity of the core clock mechanism. To this end, we repeatedly fitted the same model to available tissue-specific data [20]. Then we identified which combinations of loops generate rhythms in each tissue. Surprisingly, we found that combinations of loops which create rhythms in synergy differ depending on the tissue. The largest difference was apparent between central clocks in the brain and peripheral clocks in the rest of the body. This leads to the speculation that functional differences between organs and differences in the routes of entrainment may correspond to differing mechanisms. A variable importance of core clock components that connect to physiological functions, like nuclear receptors linking the clock with metabolism [11] in peripheral tissues, could underlie such a specificity.

One route of synchronization between clocks in different organs seem to be metabolite concentrations that cycle throughout the day [3]. Different organs rhythmically process and convert metabolites which are exchanged via the blood stream. However, the metabolism as a whole comprises a vast amount of processes, which is reflected in a high complexity of connections to the core clock as shown in a literature study. To examine the structure of the circadian metabolism on a large scale, we analyzed and integrated different types of omics data measured in liver. Rhythmic gene products and metabolites were then mapped to a comprehensive metabolic network [37] and their distribution over pathways, clustering in groups as well as temporal differences were examined. We found consistency of different omics studies and confirmed metabolism as a prevalent circadian aspect of the liver, while pointing to problems in reaction mapping. Interestingly, we identified varying average peak times for gene products mapped to different pathways. Such differences could correspond to an association of pathways with different components of the core clock [26].

1.2 Theoretical Background

The methods used for the work in this thesis are in detail descried in the respective publications and supplements in Chapter 2. Nevertheless, I want to give an overview over two basic concepts that are central for the presented work here.

1.2.1 Differential equations and rhythmic systems

Ordinary differential equations Most mathematical models in theoretical chronobiology are based on *ordinary differential equations* (ODEs) [38]. They can be used to describe the change of concentration of substances over time. For example, for a substance A

$$\frac{d[A]}{dt} = k[A]$$

describes the change of its concentration $[A]$ over time with a reaction rate proportional to it. Here, k is a constant number (the *rate constant*). This type of relation is also called *mass action kinetics*.

More generally ODEs can have the form

$$\frac{dx}{dt} = f(x,t)$$

where the function f may represent different types of kinetics, which may also be non-linear. Besides mass action kinetics, *Michaelis-Menten kinetics* [39] are often used. *Hill kinetics*, which are relevant in the scope of this thesis, can be written as:

$$f(x,t) = \frac{x^h}{K^h + x^h}$$

For x much smaller than the constant K the term is close to 0, while for x much larger than K it it close to 1 and when $x = K$, it is $1/2$. The constant h called *Hill coefficient* represents the degree of nonlinearity. That is, it determines how fast $f(x,t)$ switches from 0 to 1 when x is increased from zero to infinity.

The solution of ODEs depends on the *initial conditions*. In simulations, also referred to as numerical solutions, the changes described by the differential equations are applied starting from these values [40]. Thus, different initial conditions lead to different trajectories.

Delay-differential equations *Delay-differential equations* (DDEs) involve values of a time course that lie in the past [41]. For a substance modelled by a DDE this means that it can be affected by another substance after a certain explicit *time delay*. In a simple form DDEs can have constant discrete time delays:

$$\frac{dx(t)}{dt} = f(x(t), x(t-\tau_1), \ldots, x(t-\tau_m), t)$$

with $\tau_{1\ldots m} \geq 0$. In the context of gene regulatory networks, such delays can represent intermediate time consuming processes that would otherwise be modelled by a chain of ODEs [42].

Rhythm generation Several ODEs can be combined to an *ODE system* representing, for example, several substances that affect each others concentrations. Together, their changes define the dynamics of the system.

The following system of ODEs is a simplification of the *Goodwin oscillator* [43], which is used in Supplement S5 of the first publication in this thesis:

https://doi.org/10.18452/19960

$$\frac{dx}{dt} = \frac{1}{1 + (2 \cdot z)^h} - d_x \cdot x \tag{1.1}$$

$$\frac{dy}{dt} = x - d_y \cdot y \tag{1.2}$$

$$\frac{dz}{dt} = y - d_z \cdot z \tag{1.3}$$

For instance, x, y and z could represent mRNA, cytosolic and nuclear protein concentrations of a gene, respectively. Here, each equation has a *production* (positive) and a *degradation term* (negative). Note the nonlinear production term in equation (1.1), which might represent inhibition of the mRNA concentration by nuclear protein.

Constant values defined a priori, that is h, d_x, d_y and d_z, are called parameters. Depending on the values of the parameters the system can exhibit different behaviour. In particular, the model described in equations (1.1) to (1.3) can generate rhythms, if the Hill-coefficient is large enough. Qualitative changes like the onset of rhythms depending on parameter values are often shown in plots called *bifurcation diagrams*.

Note also, that the three equations form a *negative feedback loop*, since x activates y, y activates z and z inhibits x. In graphical representations, inhibitions are often denoted by flat arrowheads like $z \dashv x$, while activations are displayed by standard arrows $x \to y$. In fact, besides a sufficiently large nonlinearity—as given by the size of the Hill-coefficient, for example—a negative feedback is a prerequisite for rhythm generation [4, 5].

A delay in this feedback is also required and longer delays facilitate oscillation generation [5]. Therefore, in the case of DDEs one equation is sufficient to generate oscillations, where a chain of ODEs would be necessary otherwise [42].

1.2.2 Parameter optimization

In many cases parameters have to be adjusted such that a value determined by these parameters, here called *score*, is maximized or minimized. This process is called *optimization*. For each combination of n parameters, a *scoring function* (or *objective function*) may be applied to compute the score.

Together, parameters and the score create a so called *fitness landscape*—A common analogy is a landscape with hills and valleys, in which the two coordinates that determine the position on a map are the parameters and the corresponding height is the score. One could attempt to climb up as high as possible, but when reaching the top of a hill would never be sure, whether the hill was the highest one, because the landscape is large and only the local position and height are seen.

In the case of only two parameters, often a large part of the fitness landscape can be sampled. However, having many parameters leads to a combinatoric explosion and other strategies for the search of optimal scores have to be applied. To this end,

https://doi.org/10.18452/19960

a large amount of methods have been developed, with different advantages and disadvantages depending on the shape of the fitness landscape [44].

In systems biology often mathematical models are created that have parameters not directly measured in experiments. However, it may be attempted to adjust the parameters, such that simulations of the model resemble experimental observations. If a scoring function is defined that measures how well simulations and experimental measurements agree, optimization methods can be applied to estimate the parameters.

Gradient descent A commonly used method for *local optimization* involves descending iteratively along the negative gradient of a scoring function $g(x)$:

$$x_{n+1} = x_n - \gamma \nabla g(x).$$

$\nabla g(x)$ is the gradient of $g(x)$ and the step size γ can be adjusted in every iteration. The method is called local, because it only finds a local minimum, that is a parameter combination with gradient 0. However, starting from different positions in the fitness landscape different minima might be found, which could be smaller or larger.

Particle swarm optimization A method for *global optimization* used in the scope of this thesis is called *particle swarm optimization* (PSO) [45]. Different particles are distributed randomly, each representing a set of parameters. Then the particles are moved around (parameter values are changed) in different directions, readjusting their velocities in every iteration:

$$v_i(n+1) = w \cdot v_i(n) + c_1 \cdot r_1 \cdot [\hat{x}_i(n) - x_i(n)] + c_2 \cdot r_2 \cdot [g(n) - x_i(n)]$$

Here, $v_i(n)$ denotes the velocity of particle number i at step n and $x_i(n)$ its position. $\hat{x}_i(n)$ is the highest scoring position a particle has found so far, called the *individual best*, and $g(n)$ is the highest scoring position any particle has found, called the *global best*.

w is a constant called *inertia weight* that lets a particle continue moving in the direction it is currently moving. A small inertia weight results in a faster convergence of the swarm to a common location, while a high inertia weight increases exploration of the parameter space. Further, c_1 and c_2 are constants called the *cognitive* and *social component*, respectively. They determine how much particles orient towards their individual best position or the global best position found by the swarm. Finaly, r_1 and r_2 are random numbers computed in every iteration that make the swarm movement stochastic.

Once the velocities are computed, the new positions of the particles are given by:

https://doi.org/10.18452/19960

$$x_i(n+1) = x_i(n) + v_i(n+1)$$

Due to the social component particles of the swarm converge after a certain amount of iterations. Since the movement of particles is stochastic and particles may leave the local optima they have found to find better scores, this kind of method is called global optimization.

https://doi.org/10.18452/19960

Chapter 2

Publications

2.1 Published work 1: rhythm-generating loop in the circadian core clock identified

2.1.1 Context

The basic mechanism Since the noble price awarded study of a first clock protein in the fruit fly, PER [8, 46–50], a growing number of clock components has been identified [51–54]. Also in mammals, respective homologs and additional proteins are required for the generation of correctly timed circadian rhythms [2], as shown by various genetic perturbation experiments [7].

Diverse processes have been associated to mammalian clock function. For example, post-translational modifications by Casein kinase 1 [55–58] affect the period of rhythms and chromatin remodelling is linked to the binding of clock components [59–61]. However, transcriptional-translational feedback loops (TTFL) are considered the key mechanism for rhythm generation [2, 6, 7].

For generation of rhythms in *Drosophila melanogaster* a mechanism was proposed that involves PER and TIM proteins, which form a cytoplasmic dimer, translocate into the nucleus and repress their own transcription [62]. Such a mechanism was then also adopted in early mammalian models of circadian rhythm generation containing a feedback over PER/CRY as a primary mechanism instead [63–65]. As additional core clock genes were identified, models have been adapted to accomodate more components and reflect new findings [12, 23, 66–69].

A complex network In mammalian cells, CLOCK and BMAL1 form a dimer and bind to DNA sequences termed E-boxes in the promoters of their target genes, which are then activated for transcription [70]. PERs and CRYs also bind to this complex and thereby repress E-box mediated transcription [71, 72]. In addition, other DNA motifs called D-boxes and RREs (Ror-elements) are bound by further clock proteins, including DBP and REV-ERBa [9]. Each of these clock genes carries several binding boxes in its promoter, allowing other clock components to activate or repress its expression. Therefore, currently known clock genes are assembled in a tightly connected network that contains a large number of loops [9]. Figure 2.1B shows the network comprising 20 core clock genes.

https://doi.org/10.18452/19960

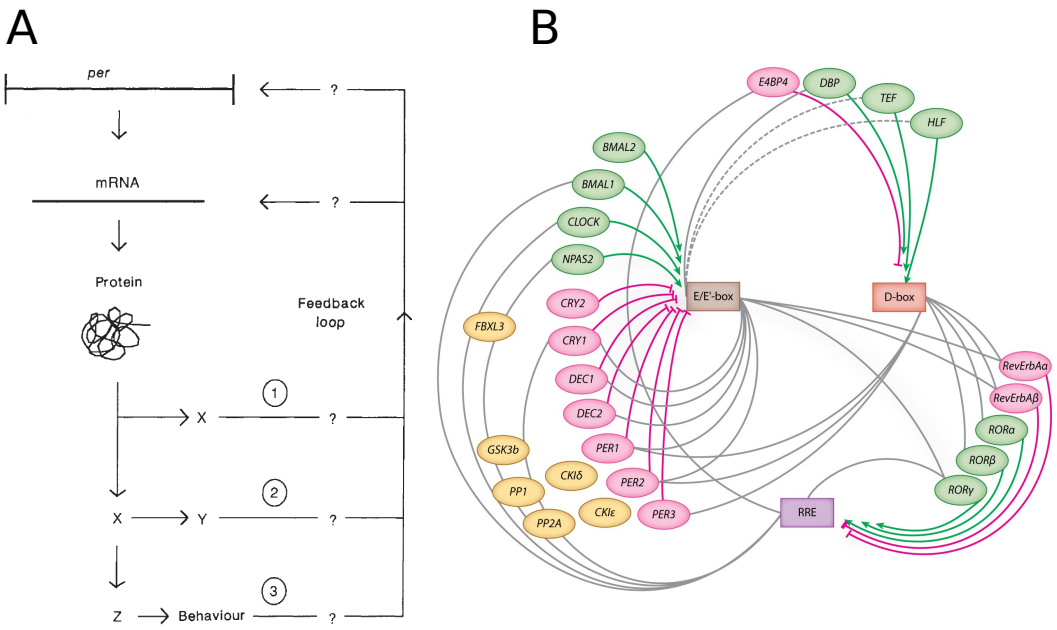

FIGURE 2.1: Complexity of the circadian core clock network. (A) Taken from [8]: first proposed mechanism of PER protein regulating its own expression. The authors were awarded the 2017 nobel prize "in physiology or medicine" for their discoveries of molecular mechanisms controlling the circadian rhythm. (B) Taken from [9]: Network circadian core clock genes after additional components had been identified. Shown genes are connected via transcriptional-translational feedback loops over cis-regulatory elements (E-boxes, D-boxes, RREs) and activate (green) or repress (pink) their targets.

With growing complexity of models, the question emerged whether the auto-inhibition of *Per* and *Cry* genes is the only key mechanism in the core clock, or whether other feedback loops could contribute to rhythm generation. A feedback comprising *Rev-erb-α* and *Bmal1* was discussed as an auxiliary loop or independent oscillator [12, 67]. Clearly, the network shown in Figure 2.1B contains many more negative feedback loops, that have not been discussed regarding their role in rhythm generation.

Unresolved molecular interactions Interestingly, *Cry1* differs from other E-box associated clock components in some aspects. Chromatin immunoprecipitation [73] and transcription data [20] show a later peak of *Cry1* expression compared to *Cry2* and *Per* genes. In addition, knockouts of *Pers* and *Crys* lead to distinct phenotypes of longer and shorter periods [74, 75] and *Cry1* knockout in contrast to *Cry2* led to a loss of rhythms in single cells [76]. Furthermore, changing the timing of expression by genetic modification of the *Cry1* promoter suggested that maintaining the late peak time of *Cry1* is also essential for circadian rhythm generation [77–79]. Based on these observations, a distinct role in rhythm generation for the E-box associated *Cry1* has been debated.

E-box mediated circadian regulation is particularly complex [73, 79–81] and the

molecular interactions are not yet completely understood. A recent study from Aryal et al. [82] gives insight into the current understanding of complex formation at E-boxes: In the active state BMAL1 and CLOCK together with additional other proteins bind to promoters in a ~0.75 MDa complex and enhance transcription. When PER1/2/3 and CRY1/2 enter the nucleus they form a much larger ~1.9 MDa complex with additional other repressor complexes including chromatin modifiers and remodellers as well as BMAL1/CLOCK, promoting its dissociation from the DNA and repression of transcription. However, exact composition of the complex and structural mechanisms regarding the conformation of proteins are not known in detail. Electron microscopy analysis suggests a dynamic and variable composition of transcriptional effectors within this quasi-spherical ~40nm complex [82].

Therefore, the complexity of interactions between circadian core clock components still remains to be fully understood, while at the same time, the physical interaction of core clock molecules is not known in mechanistic detail.

2.1.2 Research question and findings

Research question Given the complexity of the network shown in Figure 2.1B and the debated role of loops contained in this network, we asked which parts of the network are most essential for the generation of rhythms.

Different roles for modular building blocks have been described earlier for larger networks [83]. From theory it is known that a negative feedback loop is required for generating oscillations [4], but positive feedbacks might facilitate their emergence [10]. We were therefore interested, which feedback loops generate rhythms in the network.

Approach Using a data-driven mathematical model of the circadian clock [23], we systematically tested which regulations are essential by perturbing them in simulations. To this end we devised a method for analysis we call "clamping".

The model we used for analysis was published by Korenčič et al. [23]. It reflects the complexity shown in Figure 2.1B by containing all described regulations. Thus, the model is well suited to study the role of diverse feedbacks. Further, instead of including hypotheses on mechanistic interactions of molecules, it constitutes a general framework of interactions between genes, which accounts for the imprecisely characterized process of complex formation of core clock components [82]. Therefore, the model is relatively free regarding interactions of its components and well suited for an exploratory approach. For example, *Cry1* was included as a separate variable, because its expression time differs from *Cry2* and *Per* genes [20, 73].

Findings Testing all combinations of perturbations we found a network motif termed repressilator [84]—a type of negative feedback loop—that involves a combination of genes which has not been discussed as a generator of circadian rhythms before. In

the model this repressilator is both necessary, as perturbing one regulation it contains disrupts rhythmicity, and sufficient, since when perturbing all other regulations, it still produces rhythms. We compared this result with experimental evidence from the literature and discussed consistency with experimental perturbation studies. Further, we confirmed the capability of the previously considered *Rev-erb-α/Bmal1* loop [85–87] to generate 24 hour rhythms by showing that a slight change in the model parameters activates this loop.

Interestingly, the identified repressilator motif contains three groups: (i) *Per* genes and *Cry2*, (ii) *Cry1* with a distinct role within the loop and (iii) *Rev-erb* and *Ror* genes. It thus connects both loops that were discussed previously in the literature, the *Per/Cry* self-inhibition and the *Rev-erb-α/Bmal1* loop. We conclude that existence of such a loop is not an alternative option, but that several loops could act at the same time in synergy to generate robust rhythmicity.

2.1.3 Publication: Feedback loops of the mammalian circadian clock constitute repressilator

The publication [88] with `DOI:10.1371/journal.pcbi.1005266` is available at:

https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005266rev=2

2.1.4 Discussion in context

Question and approach In this publication, we approached the network complexity of the circadian core clock by analyzing a data-driven mathematical model. We devised a strategy to decompose the network into smaller parts and analyzed their behaviour, leading to the discovery of our main result that "feedback loops of the mammalian circadian clock constitute repressilator".

Naturally, our results depend on the validity of the analyzed model. A first version of the model was published in 2012 [89] and then revised and refined two years later [23]. It was carefully fitted to experimental data, taking known values of degradation rates and estimates of delays into account [23]. In addition, in the presented publication we analyzed the behaviour of the model around its default parameter set and found that bifurcations and dramatic changes of the period or amplitude do not occur close to these parameters. This confirms the robustness of the analyzed model.

Consistency of findings We then used a "clamping" strategy to inactivate certain parts of the network and examine the function of the remaining regulations. Using our clamping strategy in combination with parameter variations around the default values we identified the *Bmal1-Rev-erb-α* loop as a potential rhythm generator in

our model. This loop was not active in the default situation, but when a parameter was changed it generated self-sustained 24 hour oscillations, consistent with its discussed roles as an auxiliary loop and potential independent oscillator [2, 12].

Further, we identified a repressilator motif comprising genes which in this combination have not been discussed as a rhythm-generating mechanism: *Rev-erb-α*, *Per2* and *Cry1*. We found that this loop is responsible to generate oscillations in the model. Interestingly, the repressilator connects the previously discussed *Bmal1/Rev-erb-α* and *Per/Cry* loops. We conclude that several loops could act in synergy to generate circadian rhythms including the repressilator, which is consistent with peak phases of gene expression [20] and ChIP-seq [73] data as well as network topology [9].

While the *Bmal1-Rev-erb-α* loop was confirmed in the model, we did not find the *Per2* and *Cry1* self-inhibitions as generators of 24 hour oscillations. Instead, under some conditions the *Per2* and *Cry1* loops generated 10 or 15 hour rhythms. Such short period oscillations are explained by the single time delay associated to the genes' auto-inhibitions. Time delays were used that are consistent with literature results [23]. Even though these delays are relatively long, a sufficiently long period could not be achieved by the loops. On the other hand the delay estimates are not precise and *Per* and *Cry* auto-inhibitions could be relevant if they were higher. While we found no support for these loops, we cannot exclude them in another physiological setting.

Still, loops that consist of two or more regulations comprise several time delays that are summed up. This facilitates the generation of long period rhythms. Therefore, a more complex mechanism with a special role for *Cry1* as a late E-box inhibitor [20, 73] such as the repressilator presented in this paper might generate long period oscillations more easily.

Predicted molecular interactions The repressilator as rhythm generator would explain the late phase of *Cry1* [20, 73] as well as the importance of RRE regulation at the *Cry1* promoter [77] and disruption of rhythms by *Cry1* knockout [76]. It would further explain the loss of rhythms in *Rev-erb-α* double knockout mice [90]. These experimental observations are not explained by a simple auto-inhibition of *Per* and *Cry* genes. Rhythm generation by the repressilator also predicts an imbalance in *Per/Cry* regulations: *Rev-erb-α* would be mainly regulated by E-box inhibition via the PER complex and *Cry1* by inhibition at RREs, while *Cry1* specific inhibition would have a minor effect on *Rev-erb-α* regulation and E-box regulation via the PER complex would play a minor role for *Cry1* inhibition.

Aryal et al. [82] showed that PER and CRY proteins are present together in the nucleus only as a single complex binding to CLOCK and BMAL1 at E-boxes, with the exception of small amounts of CRY1 monomer. The authors conclude that their results provide no evidence for a separate action of CRY1, but do not exclude it either. Previously Koike et al. [73] had identified a large number of sites other than

E-boxes that are bound by each factor alone, especially for *Cry1*, including sites recognized by nuclear receptors. Interactions of *Per*s and *Cry*s with nuclear receptors have also been described before [91, 92]. Electron microscopy revealed heterogeneity in the structure of nuclear complexes, possibly due to a variable set of transcriptional effectors [82]. Even if proteins are present in one complex, differential regulation of target genes due to interactions with other regulatory elements might not be excluded. Chiou et al. [79] showed that PER regulation can work in a context-dependent manner as either repression or activation.

Taken together, the molecular mechanisms and interactions that underlie rhythm generation of the circadian clock still remain to be fully elucidated and current findings neither prove nor contradict the identified repressilator mechanism. From our theoretical analysis, which represents network structure and phase time relations, it appears as a likely candidate. Thus, the repressilator mechanism can serve as a hypothesis for experimental verification. It suggests a more detailed study of the mechanistic interactions underlying transcriptional regulation of the core clock, in particular regarding context-dependent effects as addressed by Chiou et al. [79]. Manipulation of promoters as done in constitutive expression experiments might constitute a promising approach to address such questions.

Extending the model When molecular interactions become elucidated in more detail and sufficient quantitative data is available, the mathematical model should be extended to incorporate this information. The generic and compact form of the current model reflect the lack of knowledge described above. In particular, time delays were chosen to represent time consuming intermediate processes without specifying the mechanistic details. While their use is justified in this way, they are a somewhat artificial and rough approximation from a quantitative point of view: In theory only an infinitely long chain of ODEs is equivalent to a time delayed differential equation [42]. However, in practice a smaller chain of processes might already suffice to create a delay of the length used in the model. Together with a supervised student, Elmir Mahammadov, we found out that a chain of three ODEs representing mRNA level, protein level and DNA binding (ChIP-seq) is sufficient to create the required delay of any DDE. Therefore, each DDE can be substituted by 3 ODEs and the 5-DDE model can be replaced by a 15-ODE model. Replacing the DDE- by an equivalent ODE model counters the accusation that oscillations are generated too easily with DDEs and also explains time delays biologically.

https://doi.org/10.18452/19960

2.2 Published work 2: Tissue-specific use of feedback loops for rhythm generation

2.2.1 Context

A network of clocks Circadian clocks are present in almost every cell in mammals. Their transcriptional-translational feedback loops are based on a set of core clock genes which generate 24 hour rhythms in a similar manner in diverse organs [2, 20, 89]. Different clocks are distinguished in the body, depending on the organ they are located at [3]. The brain region called Suprachiasmatic nucleus (SCN) is located where eye nerves are crossing and receives light inputs directly, synchronizing its rhythms with the day-night cycle [2, 13]. It is traditionally regarded as a master clock that synchronizes other clocks in the body [3]. In addition to central clocks located in the brain, there are peripheral clocks distributed over different organs of the body [3]. The liver for example has particularly strong circadian oscillations and is extensively studied in mice [2, 19].

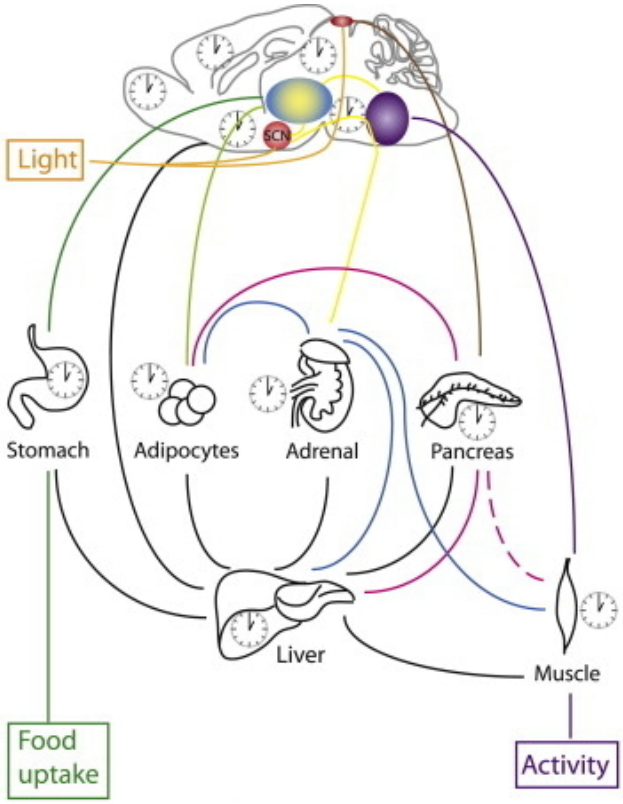

FIGURE 2.2: Network of connections between organs. Taken from [3]. Clocks are synchronized via various signals including different hormones like melatonin (brown line), ghrelin (dark green), leptin (light green), insulin/glucagon (pink) and adrenaline (blue) as well as metabolites (black lines). Further, they are linked to different external and behavioral cues such as light, feeding time and activity.

https://doi.org/10.18452/19960

Tissue-differences While the basic set of genes, expression patterns and principles generating rhythms are the same [2, 9, 20], there are differences across organs connected to rhythm generation, such as strength of coupling between cells [76]. The SCN is known to possess a particularly strong intercellular coupling [93–95], but the rhythms of separated cells are weak [96–98]. The liver on the other hand exhibits strong rhythms [20, 99], while only weak indications of coupling between hepatocytes—the most common cell type in liver—were found [100].

Also the routes of entrainment and synchronization differ naturally due to different location and type of cells. They comprise neuronal and hormonal signals as well as concentrations of cycling metabolites [3]. Furthermore, clock controlled behavior and physiology such as feeding time, activity and body temperature are sensed and feed back to the network of clocks [2]. For example, the time of food ingestion determines when metabolites in the blood stream have high concentrations and constitutes a dominant entrainment signal for many peripheral clocks [17–19].

Links to the core clock The connection between circadian clocks and metabolism is particularly complex and described in more detail in section 2.3. On the one hand rate limiting enzymes are controlled by core clock genes, on the other hand metabolites are sensed and entrain core clock gene expression [19]. Particularly notable in this context is that nuclear receptors such as the REV-ERB and ROR proteins are a part of the circadian core clock [11]. This type of transcription factor possesses both a DNA binding and a ligand binding domain, allowing to sense molecules such as hormones and metabolites. As a consequence such signalling molecules could have a differently strong effect on the core clocks of various organs, depending on the exposure of the respective cell types to these signals [101].

A large fraction of genes shows cycling expression patterns controlled by the circadian core clock [20, 102, 103], as shown by perturbation studies [102, 104]. Hence, these genes are called Clock Controlled Genes (CCGs). The sets of CCGs differ between organs, presumably reflecting different functions [20, 22, 24]. Even though less prominent then on the level of CCGs, some differences in the expression of core clock genes are observed across tissues, including phases and fold changes [20, 99]. Therefore, the question emerges, whether these differences reflect varying ways how rhythms are generated and are possibly connected to varying routes of entrainment and organ functions.

2.2.2 Research question, approach and findings

Research question Generally the same mechanism of rhythm generation is assumed for different clocks of the body [2, 3, 9]. This assumption is reasonable, given that the same genetic circuit is present and core clock genes are expressed with similar expression profiles across tissues [20, 89]. However, some differences in the expression of core clock components and varying effects of perturbations between tissues were also noted [7]. Considering the complexity of the core clock

network described in section 2.1.1, even within the same network multiple mechanisms could exist that generate rhythms. In this case expression profiles of core clock genes would not need to be dramatically different across tissues to reflect various rhythm-generating mechanisms. The question emerges, whether variations in the expression profiles of core clock genes might represent different active mechanisms, which means different feedback loops in the network.

Approach Different active feedback loops in the network correspond to different parameter sets in the model described in the previous section; keeping the same network topology, but shifting the values of parameters could change which loops are generating oscillations. We therefore used the same mathematical model [23] as in the previous study to consider further questions. The parameters determine how well a model prediction resembles the experimental expression profiles. Given inaccuracies in the experimental determination of these profiles an optimal parameter set can not be determined. However, different parameter sets can be said to fit the data well. Each tissue-specific expression profile might then be represented best by another group of parameter sets.

With the study presented in this publication we wanted to serve two goals in particular: (i) a more rigorous examination of the distribution of well-fitting parameter values for the model and (ii) examining how well-fitting parameter values vary between tissues. To this end we created a pipeline to generate ensembles of well-fitting parameter sets for each tissue. Doing so we obtained distributions of parameter sets that we then compared.

In contrast to the approach in the previous section, where just one carefully fitted parameter set chosen to represent a consensus of liver and adrenal gland data was analyzed [23, 89], we here followed a more exploratory approach. Relatively broad ranges of biologically reasonable parameter values were allowed and the distributions of well-fitting parameters were examined.

To identify the rhythm-generating feedback loops associated with a parameter set we used a clamping-strategy. This strategy is similar to the one described in the previous section, but focuses on testing feedback loops specifically. By testing combinations of loops, also synergies of multiple feedbacks can be determined.

Findings Using these methods we found that in most cases indeed multiple feedback loops generate rhythms in synergy. Looking at data from various tissues we found that the synergies differ both in the amount of involved loops as well as in their composition. In liver specific model fits the largest synergies exist and the repressilator mechanism (see previous section) is found often, which fits with its location in the network connecting several loops. In contrast, in the SCN only a few combinations of loops generate rhythms, the *Per* and *Cry* self-inhibitions supported by the *Bmal1/Rev-erb-α* loop. The *Bmal1/Rev-erb-α* loop is not found without *Per* and

Cry loops in SCN and seldomly in cerebellum fits, while other peripheral tissues like heart have a large number of fits with *Bmal1/Rev-erb-α* as the only oscillator.

2.2.3 Publication: Co-existing feedback loops generate tissue-specific circadian rhythms

The publication [105] with DOI:10.26508/lsa.201800078 is available at:

http://www.life-science-alliance.org/content/1/3/e201800078.full

2.2.4 Discussion in Context

Question and findings In this work we used gene expression data from different tissues [20], a mathematical model [23] and a fitting pipeline to address two main questions: (i) are several loops or combinations of loops able to generate rhythms that resemble data from a given tissue? and (ii) do these loops or combinations of loops differ in a tissue-dependent manner?

We found that in most cases multiple loops are combined to generate rhythms in synergy. Further, the amount of loops acting together as well as their composition depends on the tissue-specific data set. We noted three characteristic classes of synergies: In one class the *Bmal1/Rev-erb-α* loop is generally dominant and occurs as a single oscillator as well as in other combinations. This class is found for most fits to peripheral tissue data including heart, kidney and adrenal gland. Synergies found for liver fits are particularly large and include the repressilator in many cases, while single loops almost never occur. Therefore, this case corresponds to a second class. The third class is constituted by synergies found in fits to the brain tissues SCN and cerebellum. These synergies combine one or both of the *Per2* and *Cry1* loops with the *Bmal1/Rev-erb-α* loop and are therefore most restricted in the variety of combinations.

Consistency of the approach At the level of gene expression time courses [20] an earlier phase of *Cry1* and smaller amplitudes in brain tissues compared to liver for example are the most prominent differences. These are probably also deciding for the differences observed between synergies. For example, a similar phase of *Per2* and *Cry1* is expected to make it less likely for the repressilator to occur, which involves both genes as separate components. In the liver data set, where amplitudes are generally larger compared to the brain data sets, the scores of fits are on average much better. Thus, it seems to be easier for the model to generate strong rhythms with large synergies then weak oscillations with a restricted set of loop-combinations. In general, tissue-differences between experimental time courses seem to be reflected at the level of parameter values and synergies.

Our model analysis shows how different parameters associated with different rhythm-generating loops can result in structurally similar time courses showing

slight differences in amplitudes and phases. A complex network with multiple feedback loops could provide robustness and flexibility to different conditions, since rhythms might be generated in various ways. Different loops could then be emphasized in different conditions. Indeed, the conditions in organs differ in many ways, including intercellular coupling strength, dominant entrainment signals and required circadian functions [2, 3]. These differences are particularly marked between the peripheral liver and central SCN clocks, which were also associated to the most different synergies. Also on the level of parameter values the largest difference was observed between SCN and peripheral tissues.

Consistency with tissue characteristics While the strong coupling between SCN neurons in contrast to other tissues could constitute one explanation for the observed differences, the route of entrainment might be another. In SCN neurons both entrainment to light and synchronization between cells act on *Per* expression [13, 106], consistent with the *Per* loop being a characteristic feature of synergies in SCN-specific model fits. In contrast, peripheral tissues such as liver are more exposed to various metabolites which entrain clocks according to the time of feeding (since not being shielded by the blood-brain barrier) [2, 3]. The complex connection between the circadian core clock and metabolism in the liver is explained in more detail in section 2.3. Nuclear receptors such as REV-ERB and ROR proteins play an important role, since they can directly sense metabolites and hormones [11]. Interestingly, they are also an integral part of the *Bmal1/Rev-erb-α* loop, which is dominant in most fits to data from peripheral tissues. Thus, observed tissue-differences in the importance of loops that we found in our model are generally consistent with the differing dominance of entrainment routes.

In addition to tissue-dependent intercellular communication and entrainment, obviously the functions performed by organs differ. This might be reflected by the large differences observed in the expression of Clock Controlled Genes (CCGs) across tissues [20, 22, 24]. Thus, organ-specific requirements on the expression of pathways and their potential feedbacks back on the clock might also affect details of the core clock expression pattern such as phases and amplitudes.

In summary, our analysis suggests that the rhythm-generating mechanism of the core clock might be more dynamic then previously thought, assuming that multiple loops can participate in synergies. Such a system can generate oscillations in various ways depending on the context, which might constitute a general design principal.

Predictions and experimental tests A tissue-dependent mechanism would naturally imply a tissue-dependent effect of perturbations: If a loop is essential, then its disruption should lead to a loss of rhythms, while perturbations in loops that are not essential might have a minor effect. However, the ensemble based result involving distributions of synergies makes predictions less straightforward.

Nevertheless, a basic distinction between mechanisms of classes like for example brain- and peripheral tissues could be tested experimentally. Such a test would be complicated by the fact, however, that multiple loops share the same genes and often used gene knockouts are relatively invasive, since they change conditions of many other genes. In this respect a more suitable method would be constitutive expression studies and modifications of promoters that perturb only certain links in the gene network. However, while the molecular mechanisms underlying transcriptional regulation of the core clock are not fully understood [82] (see section 2.1.1) the design of such tests might be more difficult.

Therefore, direct experimental tests of our *in silico* results might not be easy at the moment, but could be attempted and are connected to the open question of better understanding links in the core clock network. Knowing which loops generate oscillations could for example be relevant to assess the consequences of mutations or drug treatments connected to molecular interactions that are specific for these loops. Apart from differences between tissues, it might be considered in the future, whether shifts in loop usage could be connected to dynamic changes of cells over time. Differences in circadian rhythms have for example been observed between SCN of young and old mice [107].

2.2.5 Presented methods

To perform our analysis we also introduced two methods, a targeted clamping strategy and vector field optimization.

Our clamping strategy is useful for testing the importance of processes in rhythmic systems. In general, an alternative version of the model is simulated in which processes are minimally changed to test a certain property. For example we test whether the property of rhythm conduction is necessary for oscillations to occur and minimally change the process by setting it to the (nonrhythmic) mean level. However, the same principle could be used in other contexts. For example to examine the response to input changes in signal transduction, with "clamped" processes being kept on a basal level. Results are more likely to be trivial in the absence of loops, though. Another way of modifying our approach could be to linearize processes instead of setting them constant, that is substituting the corresponding terms with Taylor polynomials of degree 1 instead of 0. Then the original system would be changed even less, but might still loose its rhythmicity or other nonlinear phenomena. Finally, I verified in a toy model that clamping can be used to uncover a bimodal switch that is hidden by a limit cycle. To test whether abrupt changes in a limit cycle are caused by moving between different modes of the underlying switch, a process was clamped to a constant level corresponding to the minimum on the limit cycle. Oscillations were then disrupted and a steady state was observed. The clamped value was increased up to the maximum on the limit cycle and a sudden change of the steady state was observed, corresponding to a change of the mode.

https://doi.org/10.18452/19960

Another method we introduced is vector field optimization. We used this method to improve starting conditions of a global optimization and could thereby boost the performance. The method is particularly useful when many time courses of variables are known beforehand. In cases of simple models and when the time courses used for fitting could be resembled by the model exactly, vector field optimization could even be used as a stand-alone fitting method. In Supplement S5 of the publication presented in this section we applied vector field optimization repeatedly to our circadian model and examined the distribution of the resulting starting estimates. It turned out that some of the resulting distributions are narrow compared to the allowed ranges. This might reveal information about the identifiability of parameters and importance of processes, depending on how narrow the distributions and what their absolute values are, respectively.

https://doi.org/10.18452/19960

2.3 Circadian control of liver metabolism

2.3.1 Context

The core clock and metabolism The circadian clocks present in various organs of the body are connected via the exchange of metabolites secreted into and taken up from the blood stream [2, 3]. A summary of the connections between clocks situated in different organs is shown in Figure 2.2. The circadian core clock is tightly intertwined with a large number of metabolic processes. In many cases components of the core clock direct control rate limiting metabolic reactions, while metabolites can be sensed to entrain the circadian clock [19]. The latter can happen via metabolite-activated pathways or even direct binding to nuclear receptors which are part of the core clock [11]. Taken together, the feedbacks via the regulation of reactions and metabolite concentrations back to the clock are so complex, that a simple separation of clock inputs and outputs is not possible. A summary of links between metabolism and circadian clock found in a literature search is shown in Figure 2.3. Even though the network is not complete, it demonstrates the complexity of the mutual relationship.

For example, the hormones glucagon and epinephrine activate gluconeogenesis via the recognition by G protein-coupled receptors and cAMP mediated phosphorylation of CREB (cAMP response element-binding protein), which then binds DNA motifs called cAMP response elements (CREs) and regulates transcription of the gluconeogenic key enzymes *Pck1* (phosphoenolpyruvate carboxykinase-1) and *G6pc* (glucose-6-phosphatase). Interestingly, the core clock genes *Cry1* and *Cry2* modulate the pathway by interaction with G proteins, thereby affecting transcription of these enzymes as well as blood glucose levels [108]. In another alternative pathway the CRY proteins rhythmically repress glucocorticoid receptors, which regulate the transcription of *Pck1* in response to glucocorticoids [92].

For the metabolism of lipids, for example, variants of sterol regulatory element-binding proteins (SREBPs) play an important role as transcription factors that regulate key enzymes in sterol and fatty acid synthesis. SREBP is bound to the ER membrane and tranported to golgi vesicles, where it it released by cleavage of its membrane binding domain allowing translocation to the nucleus and transcriptional regulation [109]. The release of SREBP is initiated at low sterol concentrations by SCAP and INSIG, which both sense cholesterol and oxysterol, a process that seems to be regulated by the core clock gene *Rev-erb-α* [110]. Note that this feedback, sensing low concentrations of cholesterol and inducing its synthesis, also affects other cholesterol sensitive processes such as the production of bile acid.

The synthesis of bile acid from cholesterol is regulated by nuclear receptors, notably LXR and FXR targeting the key enzyme *Cyp7a1* in a process dependent on the levels of different cholesterol derivates [111, 112]. Nuclear receptors and transcription factors involved in the regulation of bile acid synthesis are also modulated

https://doi.org/10.18452/19960

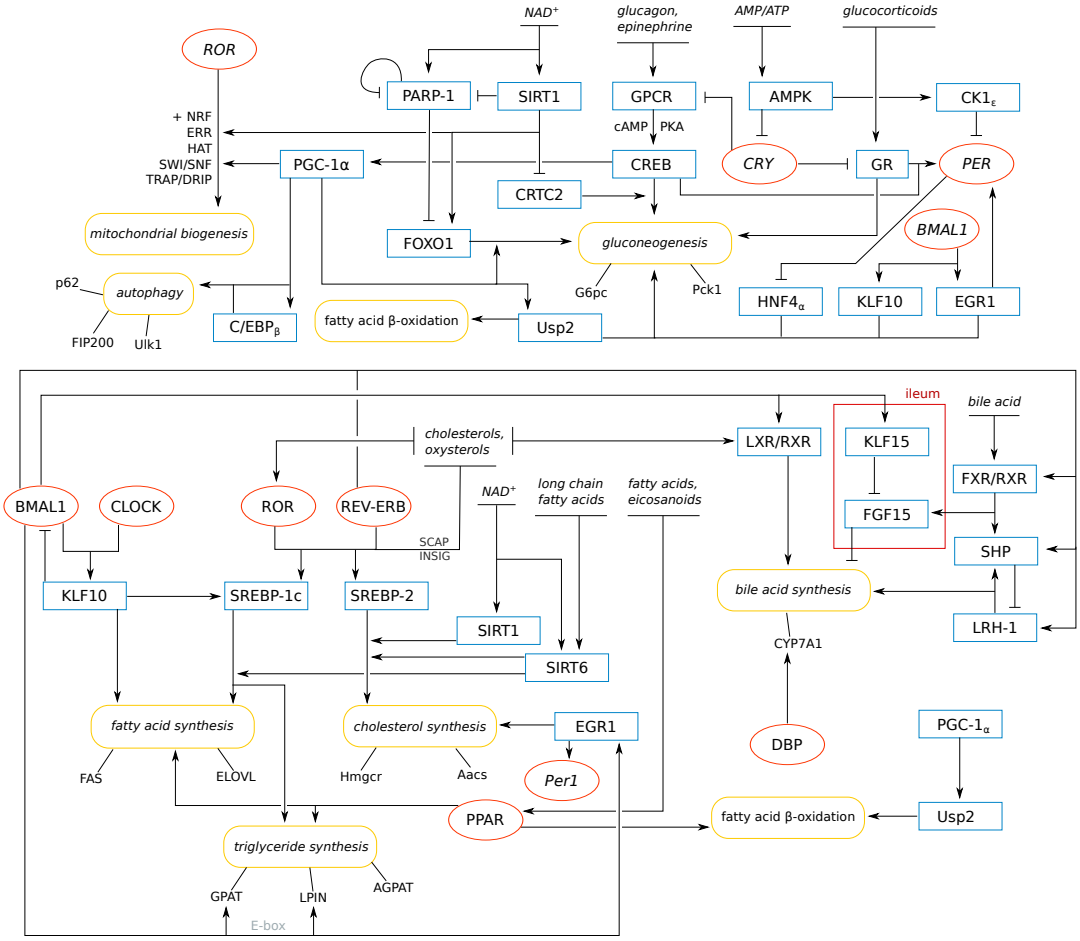

FIGURE 2.3: Mutual relation between core clock and liver metabolism. Shown are links for glucose- and lipid metabolism found in a literature study. Core clock components (red) are regulated by metabolites (underlined) via various transcription factors and co-factors (blue). On the other hand, core clock genes regulate metabolic functions (yellow) directly via rate limiting enzymes (connected by dotted lines).

by the core clock [110, 113]. Interestingly, another route regulating bile acid synthesis seems to happen via the clock controlled protein KLF15 in the ileum (a part of the small intestine) where excreted bile acid is sensed and signalled back to the liver [114].

Apart from modulating metabolic regulation pathways, the core clock is also synchronized by metabolic Zeitgeber signals. In particular, the status of energetic and metabolic levels is relayed to the core clock genes, for example via AMPK [115], SIRT [116–118], PARP-1 [119] and PGC-1α [76].

Metabolic rhythms in mouse liver The liver is a main site of metabolism involving the conversion of nutrients and storage with many processes following daily rhythms [19, 120]. It receives nutrient-rich blood directly from the portal vein and filters it through a repeated structure of hexagonal subunits made up mainly by

hepatocytes—the most common cell type in the liver. In a process called detoxification, which also takes place in the liver, xenobiotics are processed for secretion. Furthermore, bile acids are produced in the liver and secreted into the duodenum to support the digestion of lipids. The liver also exhibits strong circadian rhythms and has been extensively studied in mice as a mammalian model organism [19]. In the last two decades, also many omics data sets have been published in circadian resolution, ranging from nascent-seq [121] and transcriptomics [20, 24, 99] to ribosome profiling [122, 123], proteomics [21, 124, 125], phospho-proteomics [21, 126], ChIP-seq [73] and metabolomics [127, 128].

Many transcriptomics studies have also mapped circadian genes to metabolic key enzymes and their regulators [24, 25, 127, 129], highlighting the strong connection between circadian control and metabolism. Furthermore, proteomics and phosphoproteomics data also suggests that many rhythmic proteins with circadian periods are involved in the metabolism [21, 124–126]. While many links of the core clock to metabolic reactions and pathways have been identified, the dynamic orchestration of circadian liver metabolism as a whole is not yet fully understood.

2.3.2 Research question, approach and findings

Research question Since the metabolism appears to play an important role for clocks of the body, in particular for the synchronization of peripheral oscillators [2, 3], it is of chronobiological interest how the connection between these two aspects of physiology is established. Numerous studies have led to the emerging picture of a complex mutual relationship with many feedbacks, that makes it difficult to distinguish which are the controlling and controlled elements (see previous subsection 2.3.1). The metabolism is very complex and comprises many different processes and pathways. A general overview and distinction between its circadian aspects might be achieved by integrating and analyzing different types of omics data. We approach such an integration and mapping in the unpublished and ongoing project described here. In particular, we focus on the integration from a metabolic network perspective, to put mapped rhythmic gene products into relation. To this end, their distribution in the network should be analyzed and related to different pathways. It would be further interesting to examine how different metabolic functions can be performed under circadian constraints imposed by the data.

Approach To analyze data sets, map them to the network and examine their relation we created a bioinformatics pipeline with a workflow engine called Snakemake [130]. Different omics data sets, currently including transcriptomics [20, 99], proteomics [21, 124] and metabolomics [128], were analyzed for circadian rhythmicity with two tools in parallel: Arser [131] is a method that detects rhythms using both the time- and the frequency domain representation of a time series. A second algorithm, Rain [132], detects rhythms with a non-parametric approach that recognizes alternating up and down trends. Thus, both methods complement each other

with different perspectives on rhythmicity. To analyze the consistency of the different omics studies, in particular for the purpose of integration, their phase differences were examined.

Rhythmic gene products and metabolites were then mapped to a comprehensive metabolic network called Recon3D [37]. The network was manually curated and consistency checks of metabolic functions have been performed using Flux Balance Analysis [133]. To identify accumulations of gene products closely connected in the network, we performed a clustering. Finally, clusters were related to different metabolic pathways and their phase distributions were examined.

Findings We detected many rhythmic transcripts, proteins and metabolites, consistent with the original publications of their respective omics studies. Further, we found that a majority of metabolic reactions for which a mapping was possible was covered by rhythmic gene products. Between transcriptomics studies we found a good consistency of phases, while protein phases were about 6 hours apart from transcript phases, as also expected. Comparison of the peak times of metabolites and gene products corresponding to adjacent reactions in the network showed mostly differences of 0-6 hours. Performing a clustering analysis, we found closely connected reactions with associated rhythmic gene products that correspond to various pathways. Interestingly, different clusters and pathways showed different average phases.

This project is not finished and improvements on several aspects are expected and discussed in section 2.3.4. Nevertheless, an outline of the preliminary results is given in the following subsection.

2.3.3 Results of unpublished work

Majority of reactions associated with rhythmic data To identify rhythmic reactions we analyzed circadian time courses of different data sets [20, 21, 99, 124, 128] with two methods for rhythm detection: Arser [131], a method that uses both time- and frequency domain of the time series and Rain [132], a non-parametric method that can recognize alternations of increase and decrease. Both methods were used to detect rhythms with a 24 hour period and the intersection of their results was used for further analyses. Prior to rhythm detection the transcriptomics and metabolomics data sets were filtered for time courses with a fold change of at least 1.2 and the nuclear proteomics data set with a fold change of 1.1. Multiple testing correction was applied to all pre-filtered data sets after rhythm detection.

Transcriptomics studies have been performed often [20, 24, 99, 104] and circadian rhythms of many transcripts are clearly visible by eye in high resolution studies. In contrast, circadian proteomics and metabolomics time series, were more seldomly measured and rhythms are also harder to detect [132]. We therefore use different corrected p-values cutoffs of 0.01 and 0.2 for the two groups, respectively. Still less

rhythmic proteins and metabolites are covered and detected than genes (see Figure 2.4A). The rhythm detection setting and parameters were also tested on synthetic time series alongside the experimental data.

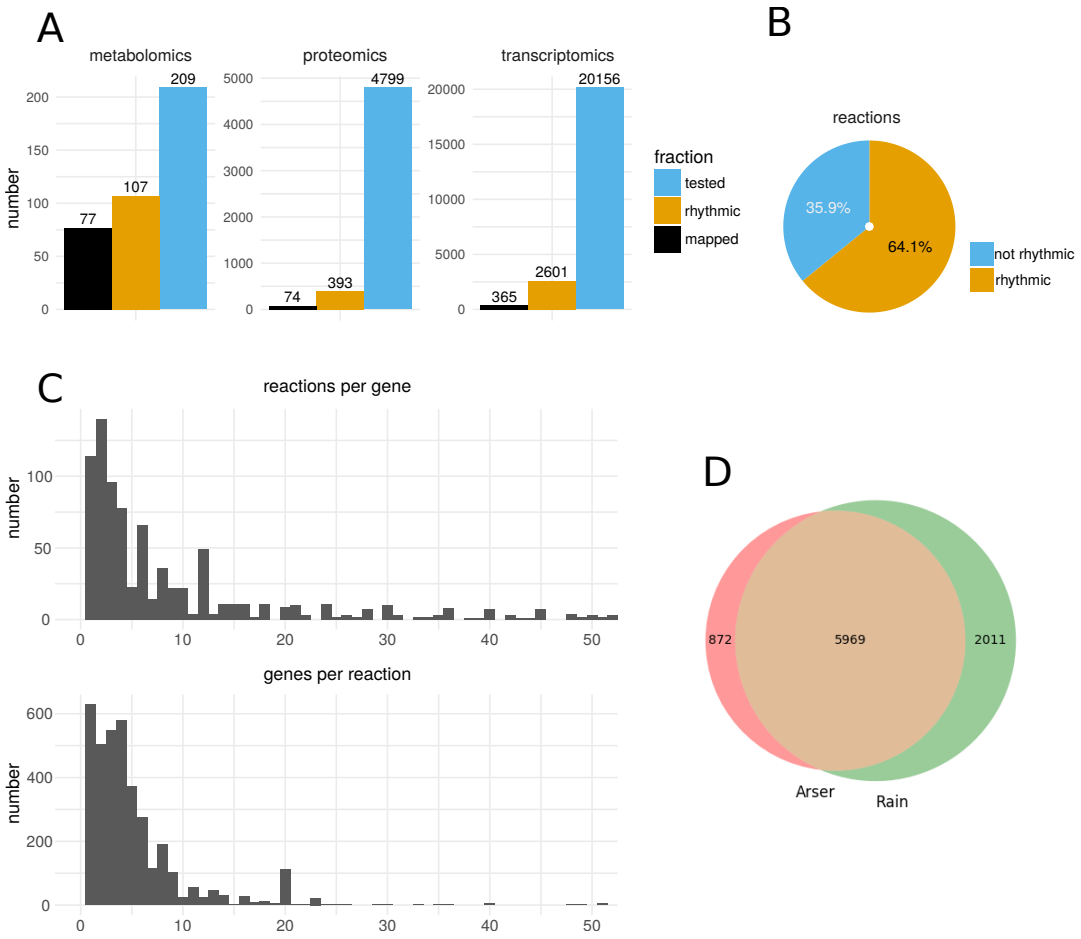

FIGURE 2.4: Detected circadian rhythms omics studies of mouse liver. (A) Numbers of tested, rhythmic and mapped time series for three data sets from [128], [124] and [20]. (B) Fractions of metabolic reactions with relevant annotation that are associated to rhythmic and non-rhythmic gene products. (C) Relation of rhythmic reactions and genes; top: histogram of how many reactions are associated to a rhythmic gene product, bottom: how many rhythmic genes products are associated to a reaction. (D) Overlaps of rhythm detection results from Arser and Rain. The overlaps were generally large. A Venn diagram of detected rhythmic probes for the data set from [99] is shown.

Consistent with the original publications [20, 124, 128] we find that large fractions of genes, proteins and metabolites (about 12%, 8% and 51%, respectively) are rhythmic. To assess their participation in metabolic reactions, we map these data sets to a comprehensive manually curated metabolic network called Recon3D [37]. This network includes a large number of reactions and metabolites, including annotations with gene products. Of the reactions that are annotated with gene information relevant for mapping, a majority is associated to at least one rhythmic gene product (see Figure 2.4B). Note however, that the relation between reactions and gene products is not unique: one gene product can participate in many reactions and one

reaction may be associated to many gene products (see Figure 2.4C). The numbers of mapped gene products shown in Figure 2.4A are examples of three representative omics-studies. In total 643 genes and 246 proteins are mapped to the metabolic network, corresponding to 3704 potentially circadian reactions.

We compare the numbers of tested reactions associated with either rhythmic or non-rhythmic gene products between pathways and find that only in 10% of the pathways less than 90% of the reactions were tested. Therefore, most annotated pathways were covered well by the mapped transcripts and proteins regarding the reactions they contain. However, since the relation between genes and reactions is not unique (compare Fig. 2.4C), the ratio of covered gene products per covered reaction varies.

For example in the fatty acid synthesis pathway on average only 0.19 rhythmic gene product are found per rhythmic reaction. This pathway involves repetitive additions of molecules to growing chains of fatty acids mediated by the same proteins. Thus, this finding is not surprising. Nevertheless, about 70% of the gene products tested in this pathway were rhythmic. In contrast to fatty acid synthesis, in the oxidative phosphorylation and linoleate metabolism pathways 4.78 and 4 rhythmic gene products were found per rhythmic reaction. Such pathways contain reactions that are mapped to many rhythmic gene products from different studies, often comprising transcriptomics as well as proteomics data.

In general, the more gene products are mapped, the more separate evidences are accumulated from different studies, resulting in a higher confidence that a reaction is rhythmic. On the other hand, the number of involved gene products varies between reactions and might depend on the characteristics of a pathway.

Other pathways with a high rhythmic-gene-to-reaction ratio of 2.56 and 2 are the propanoate metabolism and triacylglycerol synthesis, respectively. Large percentages of gene products are rhythmic in pathways for example comprising the citric acid cycle (82%), bile acid synthesis (79%) and cholesterol metabolism (76%). However, in the bile acid synthesis pathway only about 80% of the reactions were covered by mapped data. In the glycolysis/gluconeogenesis pathway about 62% of all mapped gene products are rhythmic and the rhythmic-gene-to-reaction ratio is about 1.2.

Consistent phase relationships of different data sets For identified rhythmic gene products we determined phases and amplitudes as features of the time series with unique regression fits of rhythmic models. Consistency of different data sets was then checked by inspecting the distribution of phase differences (see Figure 2.5).

All tested transcriptomics data sets [20, 124, 134] are consistent regarding their phases, with mean phase differences of less than 1 h and standard deviations of less than 3 h (compare Figure 2.5A). Also the proteomics data sets [21, 124] were consistent with a mean phase difference of about 0.3 h, however, with a higher standard deviation of about 5.6 h.

https://doi.org/10.18452/19960

Next, we examined the relation of transcript to protein phase comparing the two proteomics studies with two transcriptomics studies (Figure 2.5B and C). In both cases the protein peaks about 6 h after the transcript, while mean phase differences of the whole cell proteomics data [124] are about 2-3 h earlier than those of the nuclear proteomics data [21]). The standard deviations are high in all cases with around 5 to 7 hours, which might be expectable given the large variety of post-transcriptional and translational processes that can affect protein synthesis speed and lifetime differentially. A mean phase difference of about 6 hours has also been observed previously [124, 125].

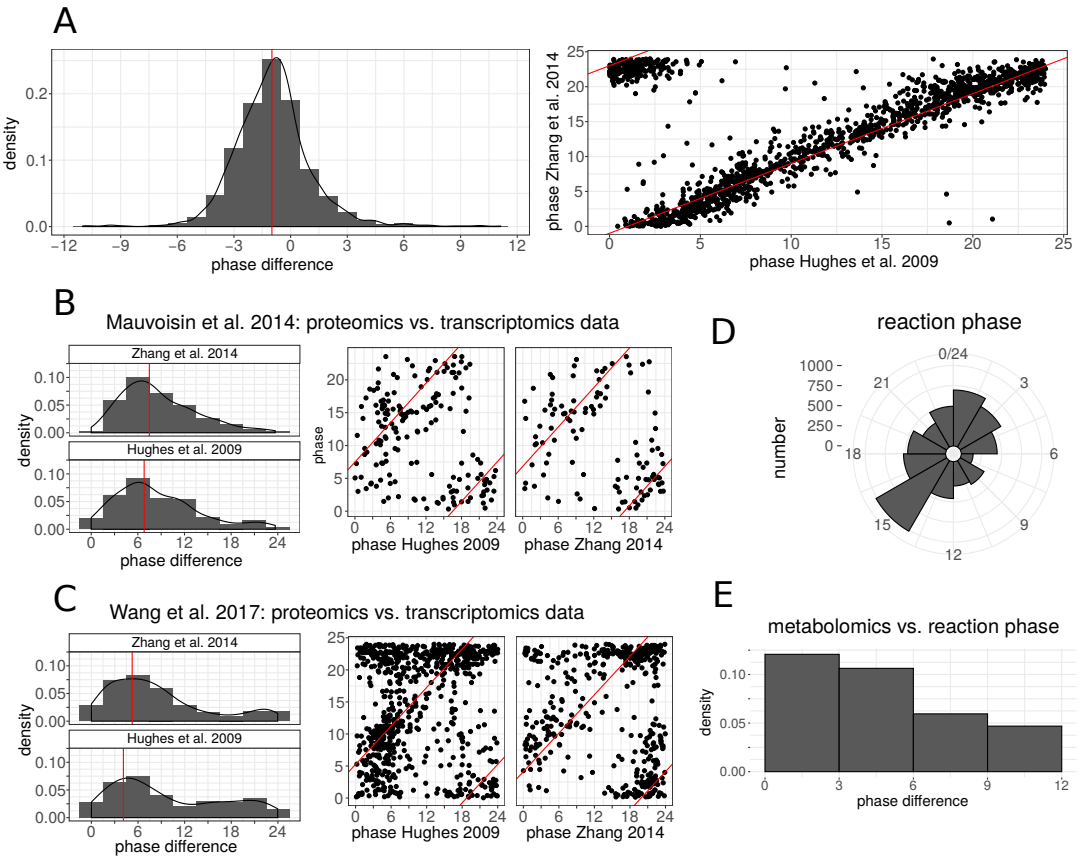

FIGURE 2.5: Relation of phases across omics studies. Shown are histograms of phase differences with a red dotted line indicating the circular mean as well as a density curve (black line) and scatter plots of phases in the corresponding two data sets with a red line of slope 1 showing the circular mean. (A) Transcriptomics data sets from [99] and [20] with a mean of -1 and sd of 1.6. (B) Proteomics data set from [124] vs. transcriptomics data set from [99] and [20] with means of 6.8, 7.5 and sd of 5.3, 4.7, respectively. (C) Nuclear proteomics data set from [21] vs. transcriptomics data set from [99] and [20] with means of 4.1, 5.3 and sd of 7, 6.2, respectively. Note that in the original publication two groups of proteins are distinguished, one that is generally correlated with transcript phase and one that is not [21]. (D) Circular histogram of predicted phases of reactions. (E) Histogram of phase differences of adjacent metabolites and reactions in the metabolic network graph.

Given these results we attempted a simple integration of transcript and protein

phases by shifting the phases of transcripts by the determined mean phase differences for the respective data sets. A phase for reactions was then determined based on the circular mean of mapped protein and shifted transcript phases. The phases of reactions are shown in Figure 2.5D. Two modes are visible at 3 h and 15 h, corresponding to 3 hours after the onset of light and dark phases respectively. Interestingly, bimodal distributions with peaks around the transitions between light and dark phases have been observed in various studies [20, 21, 124, 128]. However, on the level of integrated gene products phases are more distributed throughout the day, as expectable given the large standard deviations. Therefore, the clear modes are partly due to more reactions being associated with gene products that peak at the mode times.

Finally we also examined the relation of metabolite and reaction peak times. To this end the phases of metabolites were compared with direct neighbors in the metabolic network, which are always reactions. The phase differences of adjacent metabolites and reactions are shown in Figure 2.5E. Interestingly, more metabolites have phases similar to their associated reactions than very different ones. However, there are still many phase differences of 3-6 h, which might be consistent with delays being induced by reversibility of reactions and network motifs [135].

Clusters of rhythmic data correspond to annotated pathways To examine the distribution of mapped circadian gene products over the network we performed a clustering analysis. Clustering was done using the graph distance matrix, a matrix of pairwise distances corresponding to shortest paths in the network. In this way groups of rhythmic gene products could be found that are close to each other in the network. Since pathways can have diverse and possibly elongated shapes in the network, a single-linkage distance measure was used with agglomerative clustering [136]. The distance between clusters therefore corresponds to the shortest number of non-rhythmic reactions or metabolites separating them. Thus, clusters can be created by specifying a maximum distance that is allowed between their members. We chose a maximum distance of 3 for the results presented here.

The non-unique relation of reactions and gene products visible in Figure 2.4C implies that clustering at the level of reactions can be misleading, since several reactions of a pathway might be associated to the same rhythmic gene product. Therefore, clustering was done at the level of gene products, such that each element of a cluster represents an individual experimental evidence. A second problem concerns the connectivity of the network: Hubs, that are connected to many reactions, but are unspecific to pathways such as the metabolite H_2O, can lead to merged clusters. Therefore, hubs were removed from the network in an iterative procedure, removing those with the most connections first and recalculating connectivity. Since clusters were still merged, often by unspecific transporters, transporters were also excluded from the analysis. After removing hubs and transporters clustering was applied and

https://doi.org/10.18452/19960

the results are presented in Figure 2.6A. Clusters filtered for only those with more than 10 elements are shown.

The gene products clustered together are often annotated with the same pathways, while larger clusters comprise several connected pathways. Surprisingly, a majority of gene products associated to lysine metabolism are found in one cluster. On the other hand, a cluster comprising gene products of the citric acid cycle also includes gene products of the valine, leucine and isoleucine metabolism. Similarly, elements of the glycolysis/gluconeogenesis, pentose phosphate pathway, starch and sucrose metabolism and nucleotide interconversion are clustered together. Generally, clusters correspond to elements of connected pathways, representing groups of rhythmic metabolites and potentially rhythmic reactions that are neighboring each other in the metabolic network.

Interestingly, the phase distributions of gene products differ depending on the cluster (χ^2 test, p-value 0.008). For example, gene products associated to a cluster linked to lysine metabolism mostly peak at the end of the light phase (resting period for mice), with a p-value of 0.06 (goodness-of-fit test compared to an equal distribution). In contrast, gene products linked to the phosphatidylinositol phosphate and sphingolipid metabolism cluster mostly peak at the beginning of the resting phase (p-value 0.02), while those in a cluster associated with glycolysis/gluconeogenesis peak at the end of the active phase (p-value 0.001). Members of a cluster corresponding to both the citric acid cycle and valine, leucine and isoleucine metabolism peak not significantly during the dark/active phase (p-value 0.1), while gene products of the citric acid cycle do (p-value 0.02). On the other hand, elements of a cluster associated to inositol phosphate and glycerophospholipid metabolism have phases distributed throughout the whole day.

Independently of the cluster analysis we examined phase distributions of gene products associated to different pathways (see Figure 2.6B). Pathways comprise additional gene products that are not part of a cluster, because they are more far away from other rhythmic elements. Nevertheless, the phase distributions of many pathways and associated clusters show similar patterns. For example, gene products and metabolites participating in glycolysis/gluconeogenesis mostly peak (goodness-of-fit test, p-value 0.001) at a similar time than elements of a larger cluster also comprising several other pathways. Gene products associated to the citric acid cycle also mostly peak during the dark phase (p-value 0.02), but at slightly different times than those in the corresponding cluster. Such a peak during the active phase could correspond to a higher energy demand.

In summary, the results indicate that the peak times of gene products and metabolites differ depending on their associated larger clusters and pathways. Such a specificity at the level of pathways and clusters might reflect different functions that need to be performed at different times of the day. However, the interpretation of phase distributions still has to be substantiated by literature based expectations and more rigorous statistical tests. While they were examined for larger parts of the network,

there might also be a purpose for individual reactions with opposing phases within a pathway. For example for reactions converting metabolites in opposing directions we often observed large phase differences. Furthermore, larger phase differences might promote cycling of metabolites [135]. These questions have to be examined in more detail in the future.

2.3.4 Discussion

Idea of the approach Accumulating evidence suggests that metabolism is an important aspect of circadian physiology [90, 137]. Due to rhythms in activity and feeding, mice are challenged with nutrients and cells consume and produce metabolites at different times of the day. Circadian clocks appear to adapt to these rhythms, while at the same time controlling aspects of metabolism [2, 19]. The core clock genes are differently connected to metabolic processes [11, 26], which might have tissue-specific effects on rhythm generation as suggested in the previous section. A logical next step to understand such differences is the characterization of the temporal organization of metabolic pathways. Here, we approached to connect circadian data from different omics studies in a network.

To this end we built a pipeline for the detection of rhythms, mapping and integration of different types of data. Genes, proteins and metabolites that were mapped to a comprehensive metabolic network were then analyzed regarding their proximity of associated reactions. Further, the distribution of phases was examined showing a dependency on clusters and pathways.

Lessons from previous studies Many studies have measured circadian phases of transcriptomics data and mapped them to enzymes and regulators to infer rhythmic control of specific pathways [24, 25, 127, 129]. Furthermore, circadian cycling of many enzymes that often catalyze rate-limiting steps has been detected by proteomics studies [124, 138]. In some cases measurements of mRNA, protein and enzyme activity have suggested that enzyme activity reflects the transcription level [139]. However, it was noted that half-lives of many enzymes are long [135, 140], which would be inconsistent with circadian oscillations of protein abundance [141]. Indeed, circadian oscillations of post-translational phosphorylation are often stronger than oscillations at the proteomics level [126]. This indicates that post-translational modifications (PTMs) by regulators could constitute an important mechanism of determining metabolic enzyme activity which has not been included in our analysis yet. The inclusion of PTMs in our mapping is therefore an important next step to reflect the circadian regulation of metabolism more completely.

In addition to the before mentioned approaches mapping omics data to metabolic reactions and enzymes, a database called "CircadiOmics" has been created that integrates different types of omics data and connects them in networks [142, 143]. It allows to query genes, proteins or metabolites and displays a context of their interactions as well as circadian time courses. Our setup is similar in this respect,

although, apart from being preliminary, it does not approach the comprehensiveness of a data base. We focus instead on integrating and analyzing different omics data from the metabolic network perspective regarding their organization in groups and pathways. As a next step towards a more dynamical perspective, Flux Balance Analysis (FBA) [133] could be used to test the performance of metabolic functions under conditions of predicted enzymatic activity. For this purpose, the Recon3D network [37] constitutes a comprehensive framework in which FBA has already been performed for consistency checks of metabolic functions. It remains open, whether these functions would be performed differently at different times of the day, given phases of enzyme activity. Rhythmicity of pathways could be controlled via a few rate-limiting reactions and therefore rhythmic reaction do not necessarily have to group together in clusters. An FBA analysis would also account for such cases. Recent advances in system-wide measurements have just begun to reveal the temporal organization of metabolism, while the lack of system-wide direct enzyme activity measurements makes this characterization more difficult. Therefore, different types of data might be integrated and tested for the prediction of enzymatic activity, examining their consistency in a network context for example in clusters and with FBA simulations.

Improvements and further questions So far only transcriptomics, proteomics and metabolomics data was used, but more types of circadian omics data are available that can be used for integration including ChIP-seq [73], nascent-seq [121] and ribosome-profiling [122, 123]. As mentioned before, however, a first necessary complement would be the addition of PTM data measured in circadian omics approaches [21, 126]. To this end, regulators of enzymes should also be included in the mapping process.

We have currently not accounted for the finding, that post-transcriptional and degrading processes might play a role for differences in the rhythmicity of proteins and transcripts [124, 125, 138, 141, 144]. Further, the predictions of protein phase from mRNA are imprecise with large standard deviations. Relatively slim modes of phase distributions might result from the much larger number of detected rhythmic transcripts than proteins. Nevertheless, it might be attempted to improve the integration of transcript and protein by including additional data.

Generally, it would be interesting to further examine dynamic aspects like the rhythmic accumulation of metabolites or preservation of a steady state. For example, it was shown that rhythmic glucose release from the liver balanced out other circadian glucose variations to obtain a constant level in the blood stream [145]. This provides an example how rhythmic reactions can function to preserve a steady state. It also highlights that rhythmic supply and demand from outside a cell need to be characterized to interpret the purpose of rhythmic functions and to link organs together. For determining conditions that promote metabolic rhythms within a metabolic network, a theory was developed [135]. It suggests in particular that

differing phases of neighboring reactions may increase rhythmicity. Such findings should also be accounted for in further analyses.

Conclusion Until now, we have mapped different rhythmic gene products to a network and analyzed their co-occurrence in groups of closely connected reactions. Interestingly, different groups of reactions exhibited different predominant phases. These observations are consistent with the finding for example, that rate-limiting enzymes of glucose metabolism peak at night, while others participating in lipid metabolism peak at the daytime [124]. However, such results still have to be statistically tested in a more rigorous way and interpreted in more detail. Clearly the analysis has to be improved by considering also PTMs as drivers of enzyme activity and additional data. It should also be accounted for the finding that larger phase differences of enzyme activity are beneficial in cases when metabolic rhythms have to be generated [135].

A general overview over the circadian rhythmicity of clusters and pathways and their times of activity could serve as a starting point for more detailed analyses of individual pathways. Different pathways might also be linked to different genes of the circadian core clock [26], establishing a connection to their different phases. Further, knowing the peak times of metabolic functions relevant for the rhythmic exchange of metabolites between organs would help understanding synchronization of peripheral clocks.

https://doi.org/10.18452/19960

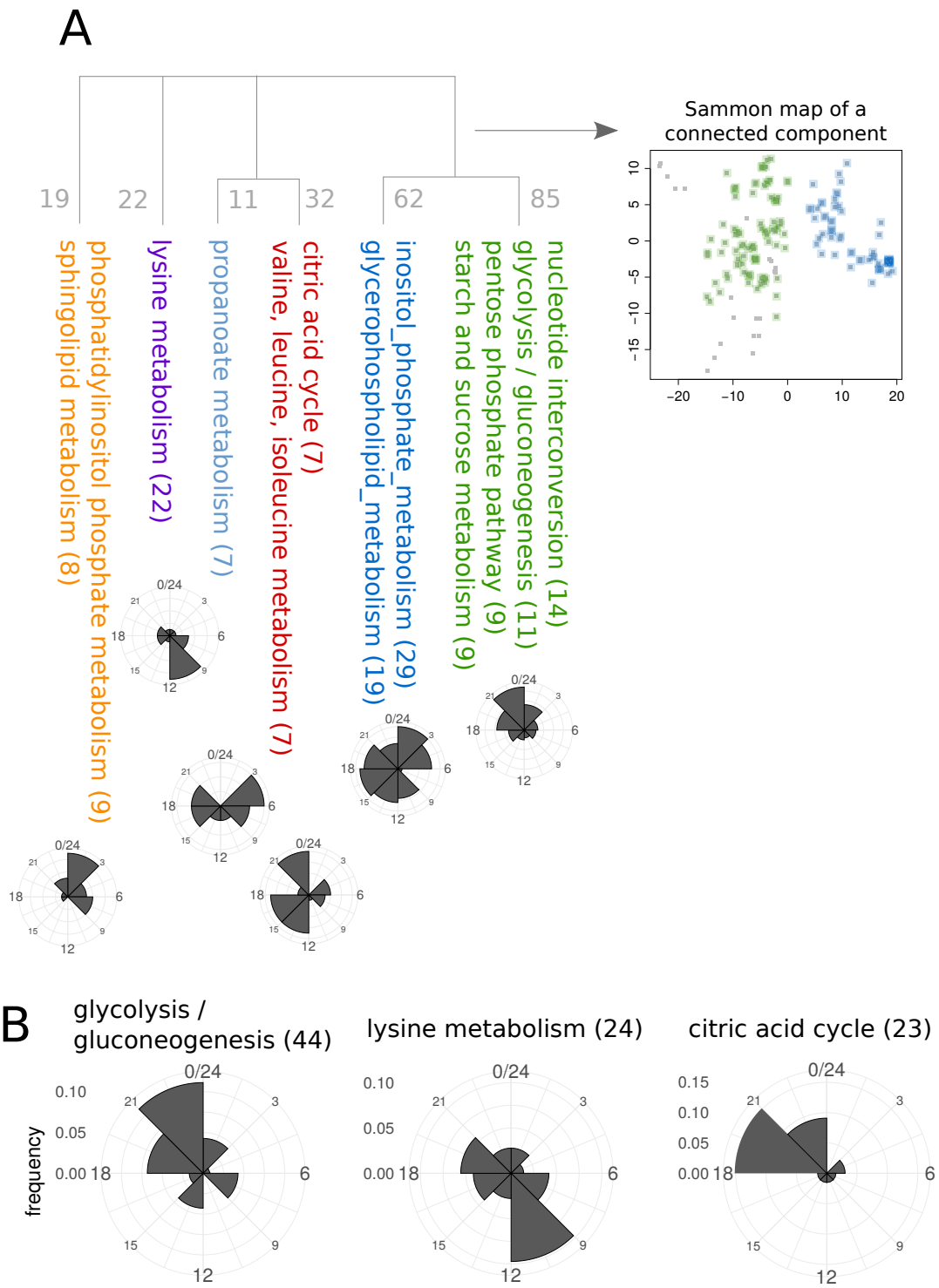

FIGURE 2.6: Phases of clusters and pathways. (A) Result of cluster analysis showing pathways that are mainly associated to the clustered gene products. The numbers of gene products are shown at the branches and clusters are distinguished by different colors. For each cluster a histogram of the phase distribution is shown below. The upper branches are created by removing hubs and transporters, the lower ones by their distance in the network. Clusters of a size < 10 have been filtered out. Right: 2D representation with multidimensional scaling showing separation of clusters on one connected component. (B) Histograms of phase distributions for gene products of three pathways (not clusters) annotated in the network.

https://doi.org/10.18452/19960

Chapter 3

General discussion and outlook

3.1 The big picture: Conclusions and outlook

With advances in molecular biology new links between genes, proteins and metabolites are discovered, leading to an emerging picture of the complexity of the involved regulations. Such regulations form networks often containing a multitude of intertwined feedbacks and the dynamics of involved components may become non-trivial. Further, there might be multiple redundant mechanisms that perform the same functions in these biological systems, possibly with a differently large contribution. Apart from predicting the behaviour of such systems, tracking down the essential regulations which form the building blocks that underlie their behaviour are thus difficult challenges. Theoretical efforts and computational tools have proven helpful for analyzing and predicting the dynamics of complex biological systems. In the first two presented papers we devised and extensively employed a strategy to combinatorically dissect and test the feedback regulations of the circadian core clock network built into a mathematical model. This strategy allowed to identify the necessary and sufficient components underlying rhythm generation in the model.

First we attempted to understand the generation of circadian rhythms in a consensus core clock model representing several tissues [23]. This data-driven model had been carefully adjusted with respect to experimental results and findings from the literature. It constitutes a compact representation of the core clock which still comprises all described feedbacks [9] without pre-emphasizing any specific mechanism by its design. Thus, the model is well suited to explore the importance of different feedback regulations based on the adjustment of model parameters to experimental data. Using our clamping strategy to combinatorically test subparts of this model, we identified the loops that can generate circadian rhythmicity. Interestingly, we found a network motif called repressilator [84] that has not been described before in this context as a necessary and sufficient component of circadian rhythm generation. A comparison with knockout studies [74, 75, 90] and additional perturbation experiments [77–79] suggests consistency of our finding. This negative feedback loop connects two other motifs that have been considered for circadian rhythms generation previously [2]. We concluded that while the repressilator constitutes a possible mechanism for circadian rhythms generation, in general a combination of

https://doi.org/10.18452/19960

several feedback loops could act in synergy.

In the second publication we applied the same successful approach to study the dependency of the core clock mechanism on tissue types. Circadian expression of genes has been measured in various mouse tissues [20]. We set up a pipeline to fit the core clock model repeatedly to the tissue-specific data sets and refine our clamping strategy to efficiently identify essential synergies of loops. Interestingly, for most tissue-specific model fits synergies of several loops were relevant for rhythm generation as previously suggested. Comparing the fitted ensembles of synergies between tissues, we found that loops have different importances in the different data sets. Differing conditions between organs may underlie the data-differences that led to different modeling predictions. We suggested that such conditions could be the functions of organs and their dominant entrainment signals that differ by tissue-type [3]. From a theoretical perspective, a network of multiple redundant loops can constitute a design principle to flexibly adapt to differing conditions, while fulfilling the same function.

A different relevance of loops and the connected core clock genes might be connected to the association of these genes with different pathways and functions [11, 26]. One of the key functions of circadian clocks seems to be the orchestration and correct alignment of metabolic processes [26, 119]. In a third project we therefore approached to characterize the circadian metabolism in more detail. Since we found in a literature study that the connections between metabolism and the core clock are as complex as metabolism itself, we approached analysis on a large scale. To this end, we analyzed several different omics data sets [20, 21, 124, 128, 134] for rhythmic components, integrated and mapped them to a large-scale metabolic network [37]. Consistency of transcriptomics and proteomics as well as metabolomics data was generally given regarding their phase times and a large fraction of the metabolic network could be associated to rhythmic data. To find accumulations of rhythmic data closely connected in the network we employed a clustering analysis that can identify groups based on their distance in the network. Interestingly, we found different predominant peak times for gene products associated to different clusters and pathways. Since core clock genes peak at different times of the day—phases are particularly spread out in peripheral tissues like the liver [20] with a strong connection to metabolism [19]—and the core clock genes seem to be linked to different metabolic pathways [26], different peak times of pathways and clusters might correspond to clock components. Understanding the correspondence of core clock and metabolism could not only help to understand possible tissue-differences in the clock mechanism, but also the synchronization of peripheral clocks which are linked via metabolite concentrations in the blood stream.

Systematically testing regulations of the core clock model adapted to specific data sets is a strategy than can easily reapplied to compare other situations. For example, metabolic gradients along lobules, the basic sub-units of the liver, seem to be

https://doi.org/10.18452/19960

connected with differential gene expression and division of functions between different zones [146]. Further, differences in circadian clock function connected to intercellular coupling have also been observed between SCN of young and old mice [107]. Apart from that, using a model that constitutes a generic framework and fitting it in an ensemble approach to different data sets, followed by a systematic model analysis might generally serve as a useful analysis pipeline. The clamping strategy we introduced constitutes a powerful tool to analyze the importance of building blocks that make up a model. A method introduced to improve the fitting step, which we called vector field optimization, might help in some cases to aid optimization.

The insights about the usage of feedback loops contained in the core clock network which we described may guide experimental exploration of the mechanistic interactions that take place at the core clock promoters. Such interactions are not fully understood [82] and combinations of different regulatory elements might lead to target gene-specific effects [79]. Further, when the connections between the circadian clock and metabolism become known in more detail, modelling of the connection [69] may be extended, also including different aspects of metabolism that are controlled by and feed back to the clock. Knowledge about the mutual relationship between clock and metabolism might also be combined with the modelling of synchronization between clocks in different organs.

An integration of our understanding of the core clock mechanism with circadian metabolism can lead to a more detailed picture of the complex synchronization of tissue-situated clocks. Further, it could help to uncover the mechanisms of metabolic disease caused by disruptions of the circadian clock [31–33] and health problems associated with a misalignment of circadian timing [34–36]. The treatment of metabolic diseases by targeting components of the core clock or associated regulators might be an option for pharmacological intervention [147].

https://doi.org/10.18452/19960

Bibliography

1. Dunlap, J. C. Molecular bases for circadian clocks. *Cell* **96,** 271–290 (1999).
2. Buhr, E. D. & Takahashi, J. S. Molecular components of the Mammalian circadian clock. *Handb Exp Pharmacol,* 3–27 (2013).
3. Albrecht, U. Timing to Perfection: The Biology of Central and Peripheral Circadian Clocks. *Neuron* **74,** 246–260 (2012).
4. Thomas, R., Thieffry, D. & Kaufman, M. Dynamical behaviour of biological regulatory networks – (I). Biological role of feedback loops and practical use of the concept of the loop-characteristic state. *Bull Math Biol* **57,** 247–276 (1995).
5. Novák, B. & Tyson, J. J. Design principles of biochemical oscillators. *Nature Reviews Molecular Cell Biology* **9,** 981–991 (2008).
6. Liu, X. *et al.* The period gene encodes a predominantly nuclear protein in adult Drosophila. *The Journal of Neuroscience* **12,** 2735–2744 (1992).
7. Ko, C. H. & Takahashi, J. S. Molecular components of the mammalian circadian clock. *Human Molecular Genetics* **15,** R271–R277 (2006).
8. Hardin, P. E., Hall, J. C. & Rosbash, M. Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. *Nature* **343,** 536–540 (1990).
9. Ukai, H. & Ueda, H. R. Systems biology of mammalian circadian clocks. *Annu Rev Physiol* **72,** 579–603 (2010).
10. Ananthasubramaniam, B. & Herzel, H. Positive feedback promotes oscillations in negative feedback loops. *PLoS One* **9,** e104761 (2014).
11. Zhao, X. *et al.* Nuclear receptors rock around the clock. *EMBO Reports* (2014).
12. Relógio, A. *et al.* Tuning the mammalian circadian clock: robust synergy of two loops. *PLoS Comput Biol* **7,** e1002309 (2011).
13. Reppert, S. M. & Weaver, D. R. Molecular Analysis of Mammalian Circadian Rhythms. *Annual Review of Physiology* (2001).
14. Balsalobre, A. *et al.* Resetting of circadian time in peripheral tissues by glucocorticoid signaling. *Science* (2000).
15. Reddy, A. B. *et al.* Glucocorticoid signaling synchronizes the liver circadian transcriptome. *Hepatology* (2007).
16. Welsh, D. K., Takahashi, J. S. & Kay, S. A. Suprachiasmatic Nucleus: Cell Autonomy and Network Properties. *Annual Review of Physiology* (2010).

https://doi.org/10.18452/19960

17. Damiola, F. *et al.* Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes and Development* (2000).

18. Stokkan, K. A., Yamazaki, S., Tei, H., Sakaki, Y. & Menaker, M. Entrainment of the circadian clock in the liver by feeding. *Science* (2001).

19. Zwighaft, Z., Reinke, H. & Asher, G. The Liver in the Eyes of a Chronobiologist. *Journal of Biological Rhythms* (2016).

20. Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc Natl Acad Sci U S A* **111,** 16219–16224 (2014).

21. Wang, J. *et al.* Nuclear Proteomics Uncovers Diurnal Regulatory Landscapes in Mouse Liver. *Cell Metabolism* (2017).

22. Storch, K. F. *et al.* Extensive and divergent circadian gene expression in liver and heart. *Nature* (2002).

23. Korenčič, A. *et al.* Timing of circadian genes in mammalian tissues. *Sci Rep* **4,** 5782 (2014).

24. Panda, S. *et al.* Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* (2002).

25. Oster, H. *et al.* The circadian rhythm of glucocorticoids is regulated by a gating mechanism residing in the adrenal cortical clock. *Cell Metabolism* (2006).

26. Brown, S. A. Circadian Metabolism: From Mechanisms to Metabolomics and Medicine. *Trends in Endocrinology and Metabolism* (2016).

27. Czeisler, C. A., Klerman, E. B., Chin, W. W. & Pfaff, D. Circadian and sleep-dependent regulation of hormone release in humans. *Recent progress in hormone research* (1999).

28. Jones, C. R. *et al.* Familial advanced sleep-phase syndrome: A short-period circadian rhythm variant in humans. *Nature Medicine* (1999).

29. Fu, L. & Lee, C. C. The circadian clock: Pacemaker and tumour suppressor. *Nature Reviews Cancer* (2003).

30. Fuhr, L., Abreu, M., Pett, P. & Relógio, A. Circadian systems biology: When time matters. *Computational and Structural Biotechnology Journal* (2015).

31. Asher, G. & Sassone-Corsi, P. Time for food: The intimate interplay between nutrition, metabolism, and the circadian clock. *Cell* (2015).

32. Milev, N. B. & Reddy, A. B. Circadian redox oscillations and metabolism. *Trends in Endocrinology and Metabolism* (2015).

33. Perelis, M., Ramsey, K. M. & Bass, J. The molecular clock as a metabolic rheostat. *Diabetes, Obesity and Metabolism* (2015).

34. Scheer, F. A. J. L., Hilton, M. F., Mantzoros, C. S. & Shea, S. A. Adverse metabolic and cardiovascular consequences of circadian misalignment. *Proceedings of the National Academy of Sciences* (2009).

35. Spiegel, K., Tasali, E., Leproult, R. & Van Cauter, E. Effects of poor and short sleep on glucose metabolism and obesity risk. *Nature Reviews Endocrinology* (2009).

36. Parsons, M. J. *et al.* Social jetlag, obesity and metabolic disorder: Investigation in a cohort study. *International Journal of Obesity* (2015).

37. Brunk, E. *et al.* Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nature Biotechnology* (2018).

38. Ingalls, B. P. Mathematical Modelling in Systems Biology : An Introduction. *Journal of Chemical Information and Modeling* (2013).

39. Michaelis, L. & Menten, M. L. Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift* (1913).

40. Butcher, J. C. *Numerical methods for ordinary differential equations* (John Wiley & Sons, 2016).

41. Driver, R. D. *Ordinary and delay differential equations* (Springer Science & Business Media, 2012).

42. Bordyugov, G., Westermark, P. O., Korencic, A., Bernard, S. & Herzel, H. Mathematical modeling in chronobiology. *Handbook of Experimental Pharmacology* (2013).

43. Goodwin, B. C. Oscillatory behavior in enzymatic control processes. *Advances in enzyme regulation* (1965).

44. Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J. A. & Blom, J. G. Systems biology: Parameter estimation for biochemical models. *FEBS Journal* (2009).

45. Poli, R., Kennedy, J. & Blackwell, T. Particle swarm optimization. An overview. *Swarm Intelligence* (2007).

46. Konopka, R. J. & Benzer, S. Clock mutants of Drosophila melanogaster. *Proc Natl Acad Sci U S A.* doi:10.1073/pnas.68.9.2112 (1971).

47. Bargiello, T. A., Jackson, F. R. & Young, M. W. Restoration of circadian behavioural rhythms by gene transfer in Drosophila. *Nature.* doi:10.1038/312752a0 (1984).

48. Bargiello, T. A. & Young, M. W. Molecular genetics of a biological clock in Drosophila. *Proc Natl Acad Sci U S A.* doi:10.1073/pnas.81.7.2142 (1984).

49. Reddy, P. *et al.* Molecular analysis of the period locus in Drosophila melanogaster and identification of a transcript involved in biological rhythms. *Cell.* doi:10.1016/0092-8674(84)90265-4 (1984).

50. Zehring, W. A. *et al.* P-element transformation with period locus DNA restores rhythmicity to mutant, arrhythmic drosophila melanogaster. *Cell.* doi:10.1016/0092-8674(84)90015-1 (1984).

51. Myers, M. P., Wager-Smith, K., Wesley, C. S., Young, M. W. & Sehgal, A. Positional cloning and sequence analysis of the Drosophila clock gene, timeless. *Science (New York, N.Y.)* doi:10.1126/science.285.5427.553 (1995).

52. Sehgal, A. *et al.* Rhythmic expression of timeless: A basis for promoting circadian cycles in period gene autoregulation. *Science.* doi:10.1126/science.270.5237.808 (1995).

53. Allada, R., White, N. E., So, W. V., Hall, J. C. & Rosbash, M. A mutant Drosophila homolog of mammalian clock disrupts circadian rhythms and transcription of period and timeless. *Cell.* doi:10.1016/S0092-8674(00)81440-3 (1998).

54. Rutila, J. E. *et al.* Cycle is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of Drosophila period and timeless. *Cell.* doi:10.1016/S0092-8674(00)81441-5 (1998).

55. Lowrey, P. L. *et al.* Positional syntenic cloning and functional characterization of the mammalian circadian mutation tau. *Science.* doi:10.1126/science.288.5465.483 (2000).

56. Akashi, M., Tsuchiya, Y., Yoshino, T. & Nishida, E. Control of intracellular dynamics of mammalian period proteins by casein kinase I epsilon (CKIepsilon) and CKIdelta in cultured cells. *Molecular and cellular biology.* doi:10.1128/mcb.22.6.1693-1703.2002 (2002).

57. Eide, E. J., Vielhaber, E. L., Hinz, W. A. & Virshup, D. M. The circadian regulatory proteins BMAL1 and cryptochromes are substrates of casein kinase I?? *Journal of Biological Chemistry.* doi:10.1074/jbc.M111466200 (2002).

58. Xu, Y. *et al.* Functional consequences of a CKIδ mutation causing familial advanced sleep phase syndrome. *Nature.* doi:10.1038/nature03453 (2005).

59. DiTacchio, L. *et al.* Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. *Science.* doi:10.1126/science.1206022 (2011).

60. Vollmers, C. *et al.* Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metabolism.* doi:10.1016/j.cmet.2012.11.004 (2012).

61. Papazyan, R., Zhang, Y. & Lazar, M. A. Genetic and epigenomic mechanisms of mammalian circadian transcription. *Nature Structural and Molecular Biology.* doi:10.1038/nsmb.3324 (2016).

62. Darlington, T. K. *et al.* Closing the circadian loop: CLOCK-induced transcription of its own inhibitors per and tim. *Science.* doi:10.1126/science.280.5369.1599 (1998).

63. Leloup, J.-C. & Goldbeter, A. Toward a detailed computational model for the mammalian circadian clock. *Proc Natl Acad Sci U S A* **100,** 7051–7056 (2003).

64. Forger, D. B. & Peskin, C. S. A detailed predictive model of the mammalian circadian clock. *Proc Natl Acad Sci U S A* **100,** 14806–14811 (2003).

65. Becker-Weimann, S., Wolf, J., Herzel, H. & Kramer, A. Modeling feedback loops of the Mammalian circadian oscillator. *Biophys J* **87,** 3023–3034 (2004).

66. Mirsky, H. P., Liu, A. C., Welsh, D. K., Kay, S. A. & Doyle 3rd, F. J. A model of the cell-autonomous mammalian circadian clock. *Proc Natl Acad Sci U S A* **106,** 11107–11112 (2009).

67. Kim, J. K. & Forger, D. B. A mechanism for robust circadian timekeeping via stoichiometric balance. *Molecular systems biology* **8,** 630 (2012).

68. Jolley, C. C., Ukai-Tadenuma, M., Perrin, D. & Ueda, H. R. A mammalian circadian clock model incorporating daytime expression elements. *Biophysical Journal.* doi:10.1016/j.bpj.2014.07.022 (2014).

69. Woller, A., Duez, H., Staels, B. & Lefranc, M. A Mathematical Model of the Liver Circadian Clock Linking Feeding and Fasting Cycles to Clock Function. *Cell Reports.* doi:10.1016/j.celrep.2016.09.060 (2016).

70. Gekakis, N. *et al.* Role of the {CLOCK} protein in the mammalian circadian mechanism. *Science* **280,** 1564–1569 (1998).

71. Sangoram, A. M. *et al.* Mammalian circadian autoregulatory loop: A timeless ortholog and mPer1 interact and negatively regulate CLOCK-BMAL1-induced transcription. *Neuron.* doi:10.1016/S0896-6273(00)80627-3 (1998).

72. Griffin Jr, E. A., Staknis, D. & Weitz, C. J. Light-independent role of {CRY}1 and {CRY}2 in the mammalian circadian clock. *Science* **286,** 768–771 (1999).

73. Koike, N. *et al.* Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338,** 349–354 (2012).

74. Van der Horst, G. T. *et al.* Mammalian {Cry}1 and {Cry}2 are essential for maintenance of circadian rhythms. *Nature* **398,** 627–630 (1999).

75. Zheng, B. *et al.* Nonredundant roles of the m{Per}1 and m{Per}2 genes in the mammalian circadian clock. *Cell* **105,** 683–694 (2001).

76. Liu, A. C. *et al.* Intercellular coupling confers robustness against mutations in the {SCN} circadian clock network. *Cell* **129,** 605–616 (2007).

77. Ukai-Tadenuma, M. *et al.* Delay in feedback repression by {Cryptochrome 1} is required for circadian clock function. *Cell* **144,** 268–281 (2011).

78. Edwards, M. D., Brancaccio, M., Chesham, J. E., Maywood, E. S. & Hastings, M. H. Rhythmic expression of cryptochrome induces the circadian clock of arrhythmic suprachiasmatic nuclei through arginine vasopressin signaling. *Proc Natl Acad Sci U S A* **113,** 2732–2737 (2016).

79. Chiou, Y.-Y. *et al.* Mammalian Period represses and de-represses transcription by displacing CLOCK–BMAL1 from promoters in a Cryptochrome-dependent manner. *Proceedings of the National Academy of Sciences.* doi:10.1073/pnas.1612917113 (2016).

80. Rey, G. *et al.* Genome-wide and phase-specific {DNA}-binding rhythms of {BMAL1} control circadian output functions in mouse liver. *PLoS Biol* **9,** e1000595 (2011).

81. Trott, A. J. & Menet, J. S. Regulation of circadian clock transcriptional output by CLOCK:BMAL1. *PLoS Genetics.* doi:10.1371/journal.pgen.1007156 (2018).

82. Aryal, R. P. *et al.* Macromolecular Assemblies of the Mammalian Circadian Clock. *Molecular Cell.* doi:10.1016/j.molcel.2017.07.017 (2017).

83. Milo, R. *et al.* Network motifs: simple building blocks of complex networks. *Science* **298,** 824–827 (2002).

84. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–338 (2000).

85. Preitner, N. *et al.* The orphan nuclear receptor {REV}-{ERB}alpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110,** 251–260 (2002).

86. Sato, T. K. *et al.* A functional genomics strategy reveals rora as a component of the mammalian circadian clock. *Neuron.* doi:10.1016/j.neuron.2004.07.018 (2004).

87. Guillaumond, F., Dardente, H., Giguère, V. & Cermakian, N. Differential control of Bmal1 circadian transcription by REV-ERB and ROR nuclear receptors. *Journal of Biological Rhythms.* doi:10.1177/0748730405277232 (2005).

88. Pett, J. P., Korenčič, A., Wesener, F., Kramer, A. & Herzel, H. Feedback Loops of the Mammalian Circadian Clock Constitute Repressilator. *PLoS Computational Biology* (2016).

89. Korenčič, A. *et al.* The interplay of cis-regulatory elements rules circadian rhythms in mouse liver. *PLoS One* **7,** e46835 (2012).

90. Cho, H. *et al.* Regulation of Circadian Behaviour and Metabolism By REV-ERB-alpha and REV-ERB-beta. *NATURE* (2012).

91. Schmutz, I., Ripperger, J. A., Baeriswyl-Aebischer, S. & Albrecht, U. The mammalian clock component PERIOD2 coordinates circadian output by interaction with nuclear receptors. *Genes and Development* (2010).

92. Lamia, K. A. *et al.* Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. *Nature* (2011).

93. Aton, S. J. & Herzog, E. D. Come together, right...now: Synchronization of rhythms in a mammalian circadian clock. *Neuron* (2005).

94. Gonze, D., Bernard, S., Waltermann, C., Kramer, A. & Herzel, H. Spontaneous synchronization of coupled circadian oscillators. *Biophys J* **89,** 120–129 (2005).

95. Bernard, S., Gonze, D., Čajavec, B., Herzel, H. & Kramer, A. Synchronization-induced rhythmicity of circadian oscillators in the suprachiasmatic nucleus. *PLoS Computational Biology* (2007).

96. Welsh, D. K., Logothetis, D. E., Meister, M. & Reppert, S. M. Individual neurons dissociated from rat suprachiasmatic nucleus express independently phased circadian firing rhythms. *Neuron* (1995).

97. Yamaguchi, S. *et al.* Synchronization of Cellular Clocks in the Suprachiasmatic Nucleus. *Science* (2003).

98. Honma, S., Nakamura, W., Shirakawa, T. & Honma, K. I. Diversity in the circadian periods of single neurons of the rat suprachiasmatic nucleus depends on nuclear structure and intrinsic period. *Neuroscience Letters* (2004).

99. Hughes, M. E. *et al.* Harmonics of circadian gene transcription in mammals. *PLoS Genetics* (2009).

100. Guenthner, C. J. *et al.* Circadian rhythms of PER2::LUC in individual primary mouse hepatocytes and cultures. *PLoS ONE* (2014).

101. Saini, C. *et al.* Real-time recording of circadian liver gene expression in freely moving mice reveals the phase-setting behavior of hepatocyte clocks. *Genes and Development* (2013).

102. Miller, B. H. *et al.* Circadian and CLOCK-controlled regulation of the mouse transcriptome and cell proliferation. *Proceedings of the National Academy of Sciences* (2007).

103. Bozek, K. *et al.* Regulation of clock-controlled genes in mammals. *PLoS ONE* (2009).

104. Vollmers, C. *et al.* Time of feeding and the intrinsic circadian clock drive rhythms in hepatic gene expression. *Proceedings of the National Academy of Sciences* (2009).

105. Kramer, A., Herzel, H., Bordyugov, G., Kondoff, M. & Pett, J. P. Co-existing feedback loops generate tissue-specific circadian rhythms. *Life Science Alliance* (2018).

106. Herzog, E. D., Hermanstyne, T., Smyllie, N. J. & Hastings, M. H. Regulating the suprachiasmatic nucleus (SCN) circadian clockwork: Interplay between cell- autonomous and circuit-level mechanisms. *Cold Spring Harbor Perspectives in Biology* (2017).

107. Ono, D., Honma, S. & Honma, K. I. Cryptochromes are critical for the development of coherent circadian rhythms in the mouse suprachiasmatic nucleus. *Nature Communications* (2013).

108. Zhang, E. E. *et al.* Cryptochrome mediates circadian regulation of cAMP signaling and hepatic gluconeogenesis. *Nature Medicine* (2010).

109. Goldstein, J. L., DeBose-Boyd, R. A. & Brown, M. S. Protein sensors for membrane sterols. *Cell* (2006).

110. Le Martelot, G. *et al.* REV-ERBα participates in circadian SREBP signaling and bile acid homeostasis. *PLoS Biology* (2009).

111. Moore, J. T., Goodwin, B., Willson, T. M. & Kliewer, S. A. Nuclear receptor regulation of genes involved in bile acid metabolism. *Crit Rev Eukaryot Gene Expr* (2002).

112. Russell, D. W. Nuclear orphan receptors control cholesterol catabolism. *Cell* (1999).

113. Lu, T. T. *et al.* Molecular basis for feedback regulation of bile acid synthesis by nuclear receptos. *Molecular Cell* (2000).

114. Han, S. *et al.* Circadian control of bile acid synthesis by a KLF15-Fgf15 axis. *Nature Communications* (2015).

115. Lamia, K. A. *et al.* AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation. *Science* (2009).

116. Asher, G. *et al.* SIRT1 Regulates Circadian Clock Gene Expression through PER2 Deacetylation. *Cell* (2008).

117. Nakahata, Y. *et al.* The NAD+-Dependent Deacetylase SIRT1 Modulates CLOCK-Mediated Chromatin Remodeling and Circadian Control. *Cell* (2008).

118. Masri, S. *et al.* Partitioning circadian transcription by SIRT6 leads to segregated control of cellular metabolism. *Cell* (2014).

119. Asher, G. *et al.* Poly(ADP-Ribose) Polymerase 1 Participates in the Phase Entrainment of Circadian Clocks to Feeding. *Cell* (2010).

120. Reinke, H. & Asher, G. Circadian Clock Control of Liver Metabolic Functions. *Gastroenterology* (2016).

121. Menet, J. S., Rodriguez, J., Abruzzi, K. C. & Rosbash, M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *eLife* (2012).

122. Janich, P., Arpat, A. B., Castelo-Szekely, V., Lopes, M. & Gatfield, D. Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames. *Genome Research* (2015).

123. Atger, F. *et al.* Circadian and feeding rhythms differentially affect rhythmic mRNA transcription and translation in mouse liver. *Proceedings of the National Academy of Sciences* (2015).

124. Mauvoisin, D. *et al.* Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver. *Proc Natl Acad Sci U S A* **111,** 167–172 (2014).

125. Robles, M. S., Cox, J. & Mann, M. In-vivo quantitative proteomics reveals a key contribution of post-transcriptional mechanisms to the circadian regulation of liver metabolism. *PLoS Genet* **10,** e1004047 (2014).

126. Robles, M. S., Humphrey, S. J. & Mann, M. Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology. *Cell Metabolism* (2017).

127. Eckel-Mahan, K. L. *et al.* Coordination of the transcriptome and metabolome by the circadian clock. *Proceedings of the National Academy of Sciences* (2012).

128. Krishnaiah, S. Y. *et al.* Clock Regulation of Metabolites Reveals Coupling between Transcription and Metabolism Cell Metabolism Clock Regulation of Metabolites Reveals Coupling between Transcription and Metabolism. *Cell Metabolism* (2017).

129. Dyar, K. A. *et al.* Muscle insulin sensitivity and glucose metabolism are controlled by the intrinsic muscle clock. *Molecular Metabolism* (2014).

130. Köster, J. & Rahmann, S. Snakemake-a scalable bioinformatics workflow engine. *Bioinformatics* (2012).

131. Yang, R. & Su, Z. Analyzing circadian expression data by harmonic regression based on autoregressive spectral estimation. *Bioinformatics* (2010).

132. Thaben, P. F. & Westermark, P. O. Detecting rhythms in time series with rain. *Journal of Biological Rhythms* (2014).

133. Varma, A. & Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Applied and Environmental Microbiology* (1994).

134. Hughes, M. E. *et al.* Brain-specific rescue of Clock reveals system-driven transcriptional rhythms in peripheral tissue. *PLoS Genet* **8,** e1002835 (2012).

135. Thurley, K. *et al.* Principles for circadian orchestration of metabolic pathways. *Proceedings of the National Academy of Sciences* (2017).

136. Everitt, B. *Cluster analysis 5th edition* (2011).

137. Green, C. B., Takahashi, J. S. & Bass, J. The Meter of Metabolism. *Cell* (2008).

138. Neufeld-Cohen, A. *et al.* Circadian control of oscillations in mitochondrial rate-limiting enzymes and nutrient utilization by PERIOD proteins. *Proceedings of the National Academy of Sciences* (2016).

139. Clarke, C. F., Fogelman, A. M. & Edwards, P. A. Diurnal rhythm of rat liver mRNAs encoding 3-hydroxy-3-methylglutaryl coenzyme A reductase. Correlation of functional and total mRNA levels with enzyme activity and protein. *Journal of Biological Chemistry* (1984).

140. Schwanhausser, B. *et al.* Global quantification of mammalian gene expression control. *Nature* (2011).

141. Lück, S., Thurley, K., Thaben, P. F. & Westermark, P. O. Rhythmic degradation explains and unifies circadian transcriptome and proteome data. *Cell Reports* (2014).

142. Patel, V. R., Eckel-Mahan, K., Sassone-Corsi, P. & Baldi, P. CircadiOmics: Integrating circadian genomics, transcriptomics, proteomics and metabolomics. *Nature Methods* (2012).

143. Ceglia, N. *et al.* CircadiOmics: Circadian omic web portal. *Nucleic Acids Research* (2018).

144. Reddy, A. B. *et al.* Circadian Orchestration of the Hepatic Proteome. *Current Biology* (2006).

145. Lamia, K. A., Storch, K.-F. & Weitz, C. J. Physiological significance of a peripheral tissue circadian clock. *Proceedings of the National Academy of Sciences* (2008).

146. Halpern, K. B. *et al.* Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. *Nature* (2017).

147. Solt, L. A. *et al.* Regulation of circadian behaviour and metabolism by synthetic REV-ERB agonists. *Nature* (2012).